AUTOMATIC BOUNDING ESTIMATION IN MODIFIED NLMS ALGORITHM

K. Shahtalebi and A. M. Doost-Hoseini

Department of Electrical and Computer Engineering, Isfahan Univ. of Tech.
Isfahan, Iran, shahtalebi@wongsaye.com alimdh@cc.iut.ac.ir

(Received: December 4, 1999 – Accepted in Final Form: August 15, 2001)

Abstract Modified Normalized Least Mean Square (MNLMS) algorithm, which is a sign form of NLMS based on set-membership (SM) theory in the class of optimal bounding ellipsoid (OBE) algorithms, requires a priori knowledge of error bounds that is unknown in most applications. In a special but popular case of measurement noise, a simple algorithm has been proposed. With some simulation examples the performance of algorithm is compared with MNLMS.

Key Words OBE, NLMS, MNLMS Algorithm, Overbounding, Underbounding, MLS Noise

چکیده یکی از شیوه های مناسب که طی دو دهه اخیر در طراحی الگوریتمهای وفقی مورد توجه قرار گرفته است، طراحی الگوریتمهای موسوم به OBE است. دانش طراح از سطح ماکزیمم نویز قابل اندازه گیری و بیضی گون در بردارنده پارامتر، مهمترین فرضیه مورد استفاده در طراحی این الگوریتمها است. در این راستا با جایگزینی فوق کره به جای بیضی گون، الگوریتمی به نام MNLMS ارائه شده است. ساختار الگوریتم بسیار ساده بوده و در رده الگوریتم NLMS و از خانواده الگوریتمهای OBE است. در این الگوریتم، به منظور غلبه بر محدودیت آگاهی از سطح ماکزیمم نویز، از یک شیوه جدید برای تخمین ماکزیمم دامنه نویز استفاده شده است. نتایج شبیه سازی، کارایی الگوریتم حاصل را در شرایط مناسب تایید می کنند.

INTRODUCTION

OBE algorithms are used to identify a real model of the general form

$$y_n = W^T X_n + v_n \tag{1}$$

in which $W^T = [w_1, \ldots, w_m]$ is the unknown parameter vector, $\{v_n\}$ is a disturbance, error, or input sequence and $\{X_n\}$ is a measurable sequence of m-vectors. It is assumed that for each n, v_n is bounded in magnitude by γ^*, i.e.

$$v_n^2 \le (\gamma^*)^2 \tag{2}$$

Equations 1 and 2 together yield

$$(y_n - W^T X)^2 \le (\gamma^*)^2 \tag{3}$$

Let S_n be a subset of R^m defined by

$$S_n = \{W : (y_n - W^T X_n)^2 \le (\gamma^*)^2, \quad W \in R^m\} \tag{4}$$

From a geometrical point of view, S_n is a convex polytope. Thus with each measured pair (y_n, X_n), Equations 1 and 2 yield a convex polytope in the parameter space. At any instant n, the intersection of the sequence of $S_1, \cdots, S_n$ contains W and so must any ellipsoid that bounds this intersection. OBE algorithms start with a sufficiently large ellipsoid that covers all possible values of W.

After (y_1, X_1) is acquired, an ellipsoid that bounds the intersection of the initial ellipsoid and S_1 is found. Every OBE algorithm uses a specific optimization criterion to find this ellipsoid that is denoted by E_1. By the same token, the algorithm obtains a sequence of optimal bounding ellipsoids $\{E_n\}$. The estimate for W at the nth instant is defined to be the center of E_n. Suppose that E_{n-1}, at instant $n-1$, is given by

$$E_{n-1} = \left\{ W : (W - W_{n-1})^T P_{n-1}^{-1} (W - W_{n-1}) \le \eta_{n-1}^2 \right\} \quad (5)$$

for some positive definite matrix P_{n-1} and a nonzero scalar η_{n-1}. Observing (y_n, X_n), an ellipsoid that bounds $E_{n-1} \cap S_n$ is given by

$$E_n = \left\{ W : (W - W_n)^T P_n^{-1} (W - W_n) \le \eta_n^2 \right\} \quad (6)$$

where

$$P_n^{-1} = (1 - \lambda_n) P_{n-1}^{-1} + \lambda_n X_n X_n^T \quad (7)$$

or equivalently (using matrix inversion lemma)

$$P_n = \frac{1}{1 - \lambda_n} \left[P_{n-1} - \frac{\lambda_n P_{n-1} X_n X_n^T P_{n-1}}{1 - \lambda_n + \lambda_n X_n^T P_{n-1} X_n} \right] \quad (8)$$

$$e_n = y_n - X_n^T W_{n-1} \quad (9)$$

$$W_n = W_{n-1} + \frac{\lambda_n P_{n-1} X_n}{1 - \lambda_n + \lambda_n X_n^T P_{n-1} X_n} e_n \quad (10)$$

$$\eta_n^2 = (1 - \lambda_n) \eta_{n-1}^2 + \lambda_n \gamma^2 - \frac{\lambda_n (1 - \lambda_n) e_n^2}{1 - \lambda_n + \lambda_n X_n^T P_{n-1} X_n} \quad (11)$$

and λ_n is any scalar in (0,1) [5].

In MNLMS algorithm, P_n is replaced by a diagonal matrix $\mu_n I > P_n$ (where A >B means A-B is positive definite) and an expanded set $\overline{E}_n$ where

$$\begin{aligned} \overline{E}_n &= \left\{ W : \mu_n^{-1} (W - W_n)^T (W - W_n) \le \eta_n^2 \right\} \\ &= \left\{ W : (W - W_n)^T (W - W_n) \le \mu_n \eta_n^2 \right\} \end{aligned} \quad (12)$$

which covers E_n .i.e.

$$E_n \subseteq \overline{E}_n \quad (13)$$

Choosing the value of λ_n which minimizes $\zeta_n^2 = \mu_n \eta_n^2$ leads to a very simple algorithm named MNLMS [12] (and also [10] for a geometrical approach).

$$W_n = \begin{cases} W_{n-1} & \|e_n\| \le \gamma^* \\ W_{n-1} + \dfrac{\|e_n\| - \gamma^*}{X_n^T X_n} X_n \operatorname{sign}(e_n) & \|e_n\| > \gamma^* \end{cases} \quad (14)$$

$$\zeta_n^2 = \begin{cases} \zeta_{n-1}^2 & \|e_n\| \le \gamma^* \\ \zeta_{n-1}^2 - \dfrac{(\|e_n\| - \gamma^*)}{X_n^T X_n} & \|e_n\| > \gamma^* \end{cases} \quad (15)$$

Although ζ_n^2 does not have any direct role in MNLMS algorithm, but helps distinguishing the variation of parameters. With the assumption that γ^* is chosen correctly and under ideal time-invariance condition, ζ_n^2 never goes negative (see [12]). Every time ζ_n^2 assumes a negative value, a variation in the true parameter has occurred.

However we focus on another important problem: MNLMS algorithm like conventional OBE algorithm [8] is based on the premise that $\{v_n\}$ has an upper bound that is known apriori, $\|v_n\| \le \gamma^*$, for all n. However since $\{v_n\}$ is unobservable, choosing a proper γ^* (or bounding sequence $\{\gamma^*\}$ for the case of time variable maximum level), is critical in practice. Over bounding increases the estimation error and leads to *inconsistent* estimator [11]. Under bounding is riskier because it can cause divergence. In the next section we focus on the case that over bounding or under bounding has been occurred and propose a method to decrease or increase γ_n to its correct value (γ^*). As stated earlier, our method is valid for a special class of measurement noises, which is defined in the next section.

MLS NOISE AND PROPOSED ALGORITHM

Definition 1 $\{v_n\}$ is called a Maximum Level Selecting (MLS) noise of order N if for any set of time instants $n_0, n_0+1, \cdots, n_0+N-1$, there exists at least one k such that

$$\|v_k\| = \gamma^* \quad \text{with} \quad \text{Prob.} 1$$

where γ^* is the global maximum magnitude of $\{v_n\}$. i.e.

$$\|v_n\| \le \gamma^*$$

This class, choosing a suitable N, encompasses a broad variety of noises, e.g. on-off, hard limited and quantizer systems noises. The following theorem is the basic key for noise bound correction and completing of MNLMS algorithm.

Theorem 1 Suppose $\{v_n\}$ is an i.i.d MLS noise of order N with $\|v_n\| \le \gamma^*$ and $\{u_n\}$ is an i.i.d sequence for which v_n is independent from u_n for all n. Then for every $n_0, 0 < \gamma \le \gamma^* and\ \varepsilon > 0$, there exists a positive number M such that for every K ≥ M

$$P\{\|v_n + u_n\| < \gamma \quad n = n_0, n_0+1, \cdots, n_0+K-1\} < \varepsilon \tag{16}$$

Proof*:* See the appendix

Now suppose $\{v_n\}$ is an MLS noise of order N and parameter γ^* and for a period M>>N the sequence $\{e_n\}$ in Equation 9 satisfies

$$\|e_n\| = \|v_n + X_n^T(W - W_{n-1})\| < \gamma$$

$$n = n_0, n_0+1, \cdots, n_0+M-1 \tag{17}$$

According to theorem 1 for $\gamma \le \gamma^*$ and a sufficiently large M, the probability of the above event is approximately zero. So it is clearly found that with a high degree of accuracy

$$\gamma^* < \gamma$$

Hence

$$\|v_n\| \le \gamma^* < \gamma \tag{18}$$

The above statement is based on the assumption that $\{u_n\} = \{X_n^T(W - W_{n-1})\}$ is an i.i.d sequence and u_n independent of v_n for all n (ordinary assumptions in the literature of adaptive algorithms). So we can candidate $\gamma^* = \gamma - \delta$ for the maximum noise level. i.e.

$$\gamma \to \gamma - \delta$$

where δ is an arbitrary small positive value. On the other hand, because of the nature of OBE algorithm, they use only a few percentages of the input data. So if the algorithm uses input data successively for a period exceeding L (usually L = 1, 2 or 3) without interruption, it insures that γ is less than γ^*. Hence we should increase γ

$$\gamma \to \gamma + \delta$$

The above explanation is the foundation of new algorithm called Automatic Bound Estimation (ABE) MNLMS algorithm and is summarized as follows (See [9] for another approach to ABE algorithms):

Initialization: Set $W_0 = 0, \zeta_0 = a$ (large) positive number, $\gamma_0 =$ any (over) estimated bound, Choose δ (small positive number), M (M>>N) and L (usually L = 1, 2 or 3)

$$\text{if } \|e_n\| < \gamma_n : \begin{cases} \gamma_n = \gamma_{n-1} \\ k = k+1, l = 0 \\ W_n = W_{n-1} \\ \zeta_n^2 = \zeta_{n-1}^2 \\ \text{if } k > M : \quad \gamma_n = \gamma_{n-1} - \delta \quad , k = 0 \end{cases}$$

$$\text{if } \|e_n\| > \gamma_n : \begin{cases} W_n = W_{n-1} + \dfrac{\|e_n\| - \gamma_n}{X_n^T X_n} X_n \operatorname{sign}(e_n) \\ l = l+1 \quad , k = 0 \\ \zeta_n^2 = \zeta_{n-1}^2 - \dfrac{(\|e_n\| - \gamma)^2}{X_n^T X_n} \\ \text{if } l > L : \quad \gamma_n = \gamma_{n-1} + \delta , \quad l = 0 \end{cases}$$

$$\text{if } \zeta_n^2 \le 0 : \zeta_n^2 = a\,(\text{large}) \text{ positive number}$$

With a rather tedious mathematical analysis one can show that if M, $\{v_n\}$ and $\{u_n\} = \{X_n^T (W - W_{n-1})\}$ satisfy the conditions of theorem 1, γ_n finally settles in the interval $[\gamma^* - \delta, \gamma^* + \delta]$. We skip the exact proof but demonstrate this fact via computer simulations in the following section.

SIMULATION

In this section we present simulations that support the abilities of proposed ABE-MNLMS algorithm. We compare the results with those of MNLMS algorithm using an AR(2) model with

$$y_n = cy_{n-1} + dy_{n-2} + v_n = X_n^T W + v_n \tag{19}$$

where

$$W = [c \quad d]^T \quad , \quad X_n = [y_{n-1} \quad y_{n-2}]^T$$

Four cases are considered:

Case 1. Time Invariant Parameter with Colored Noise Using $c = 1$ and $d = -0.5$ and v_n is a colored non-zero mean error sequence generated by a correlated sequence $\{w_n\}$

$$v_n = \begin{cases} 1 & \text{if } \; w_n > -1 \\ -1 & \text{otherwise} \end{cases} \tag{20}$$

in which w_n is generated by

$$w_n = -0.8 w_{n-1} + z_n$$

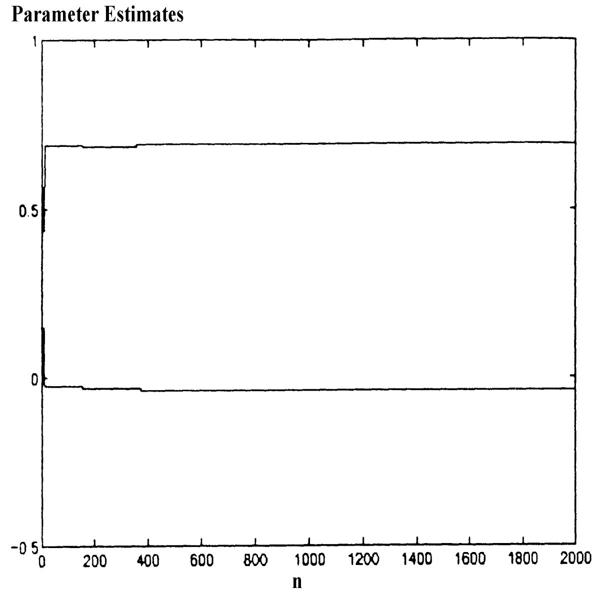
where z_n is i.i.d uniform in [-1,1]. Both algorithms are run with an overestimated bound $\gamma = 1.5$ since the true error bound $(\gamma = 1)$ is assumed unknown. The results are shown in Figure 1 (See also the result of SM--SA OBE algorithm used in [13]).

Case 2. Time Invariant Parameter with Multi Level Noise In this case v_n assumes values {-1,-2/3,-1/3,0,1/3,2/3,1} with equal probabilites. Other conditions are the same as case 1. The results are shown in Figure 2.

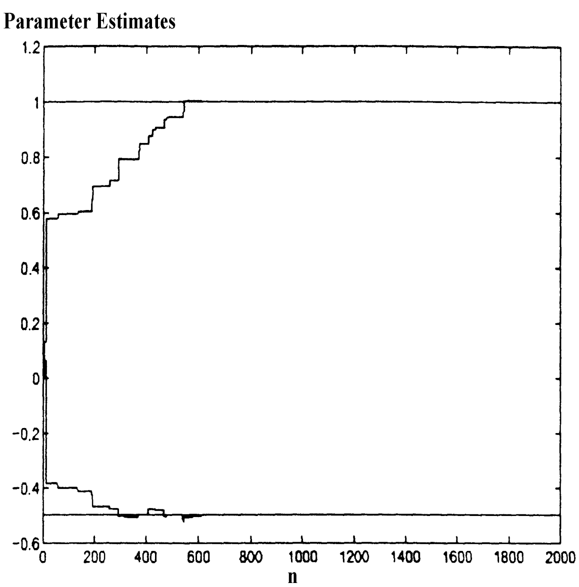
Case 3. Time Varying Model The parameter c was changed by 50 % at one-thousandth sample, while C was kept constant at its nominal value. As before, v_n chooses values {-1,-2/3,-1/3,0,1/3,2/3,1} uniformly. The parameter estimates are plotted against the true values in Figure 3. The proposed algorithm also has remarkable performance for the case of under bounding.

Case 4. Time Varying Model with Under Bounding Initial Value Consider case 3 but with $\gamma = 0.8$. The results are illustrated in Figure 4.

ABE-MNLMS algorithm exhibits improved performance over MNLMS in all cases. For example in opposition to MNLMS algorithm, the steady state error of the parameter estimate in ABE--MNLMS algorithm is zero in all cases. Especially from Figure 4 in the case of under

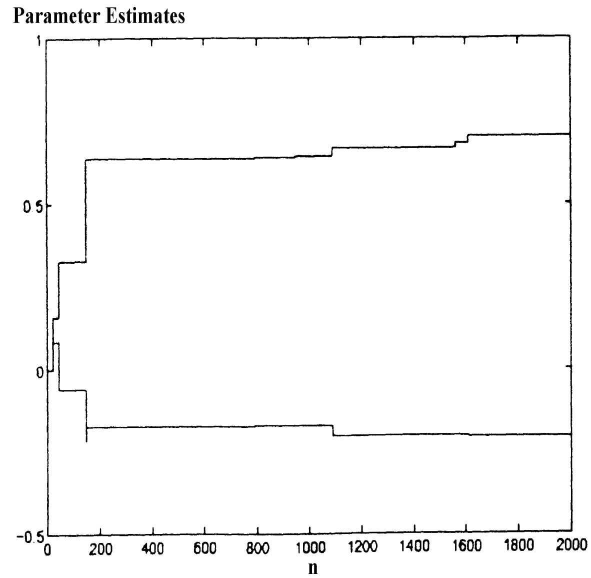


(a)

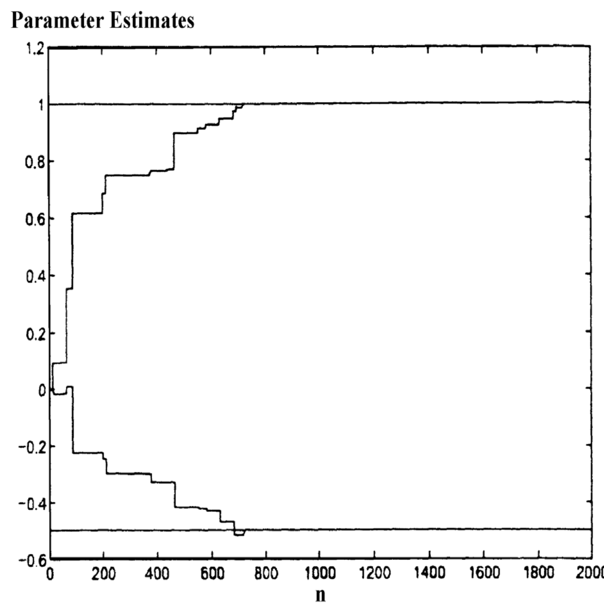


(b)

Figure 1. Parameter estimates for case 1 (a) MNLMS with $\gamma = 1.5$ and (b) proposed algorithm with $\gamma = 1.5$, M = 50 and L = 2.



(a)



(b)

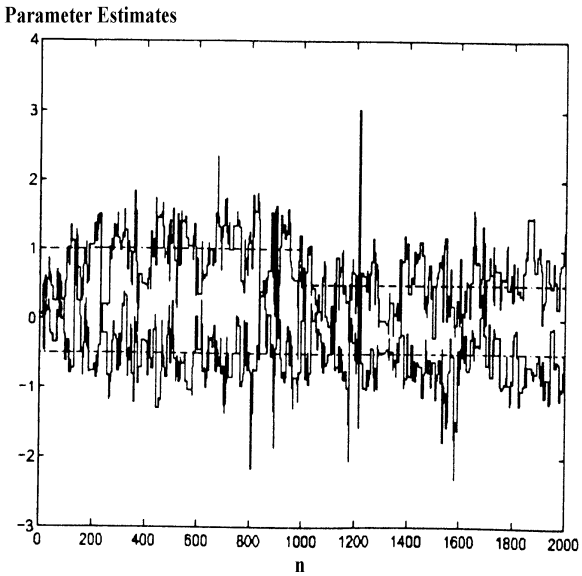
Figure 2. Parameter estimates for case 2 (a) MNLMS with $\gamma = 1.5$ and (b) proposed algorithm with $\gamma = 1.5$, M = 50 and L = 2.

bounding, MNLMS is unstable. Despite of all results, it is important to point out that MNLMS algorithm has remarkable performance when true γ is available [12].

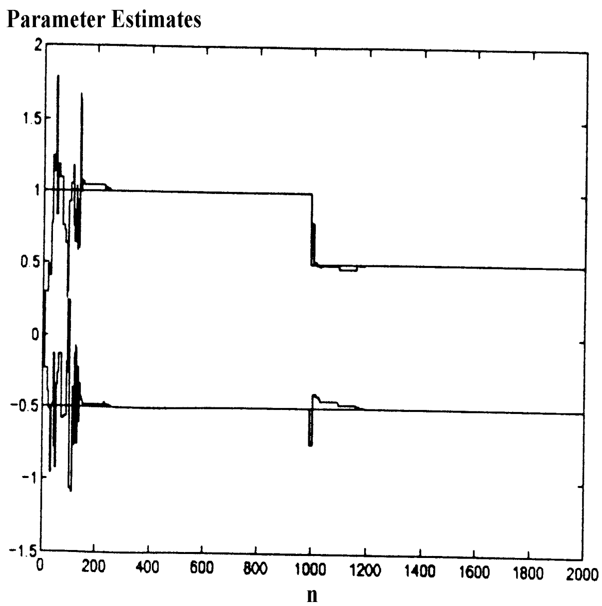
As mentioned in the end of section 2, γ_n finally settles in the interval

$$[\gamma^* - \delta, \gamma^* + \delta]$$

To illustrate this fact, Figure 5 shows the

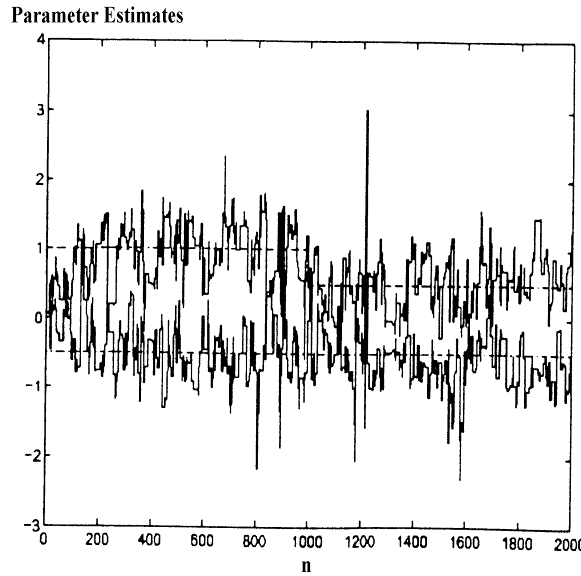


(a)

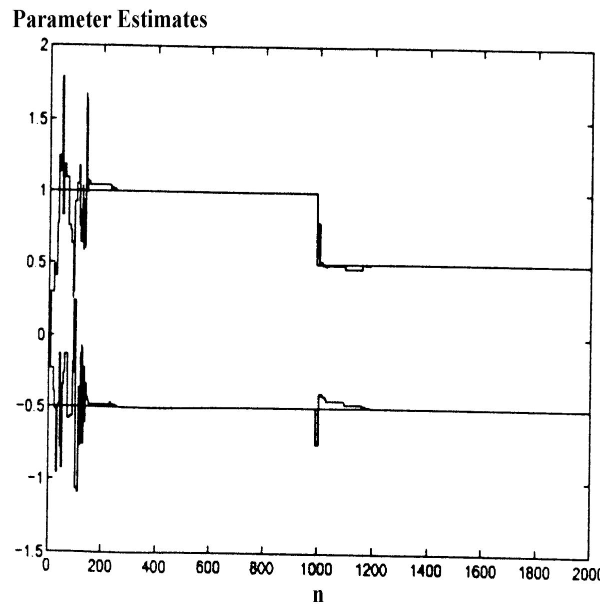


(b)

Figure 3. Parameter estimates for *case 3* (a) MNLMS with $\gamma = 1.5$ and (b) proposed algorithm with $\gamma = 1.5$, M = 50 and L=2.



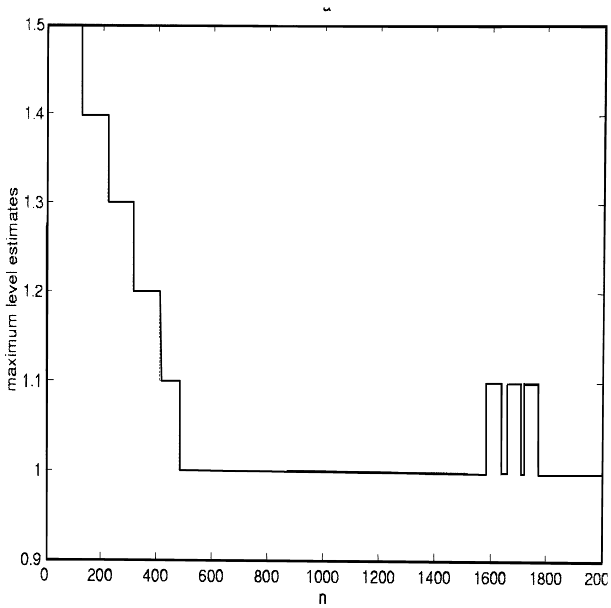
(a)



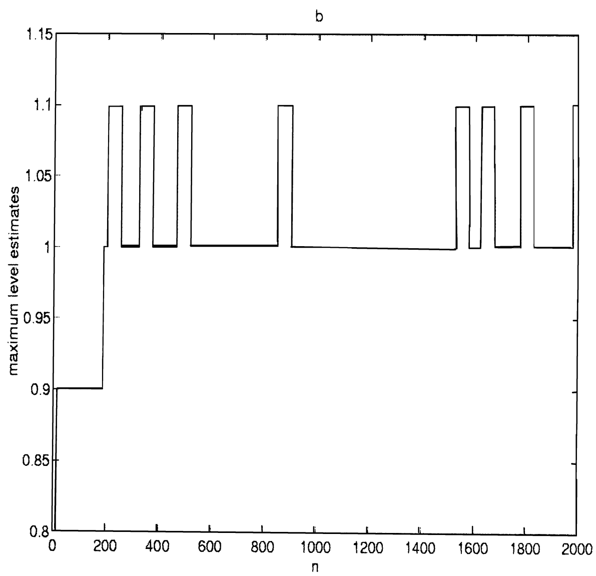
(b)

Figure 4. Parameter estimates for *case 4* (a) MNLMS with $\gamma = .8$ and (b) proposed algorithm with $\gamma_0 = 1.5$, M = 50 and L=2.

estimated values of $\gamma^* \left(i.e.\, \gamma_n\right)$, calculated by ABE--MNLMS algorithm in the cases 3 and 4. In the cases 3 and 4 , the initial value of γ was $\gamma_0 = 1.5$ and $\gamma_0 = 0.8$ respectively. Figure 5 (a and b) shows that after $n = 500$ (in case 3) and $n = 200$ (in case 4) ABE--MNLMS algorithm has found its true value. Because of the value of L (L=2) that is small, there is not any underestimating after these time instants.



(a)



(b)

Figure 5. Estimates of γ* in (a) Case *3* (over bounding) and (b) Case 4 (under bounding).

CONCLUSION

A simple strategy to find the true maximum level of noise has been derived. It is valid for those measurement noises that reach their maximum level in finite durations with probability one. Simulation results show that the tracking performance of this algorithm in finding true maximum level of noise is comparable to that of MNLMS algorithm.

APPENDIX: PROOF OF THEOREM 1

Because $\{v_n\}$ is an MLS noise of order N, there are time instants $\gamma \le \gamma^*$ in the sets of length N such that

$$v_{k_i} = \gamma^*,\ k_i \in \{n_0 + iN, n_0 + (i+1)N - 1\},\ i = 0,1,2,\ldots \tag{21}$$

Suppose K is an integer multiple of N. Because $\gamma \le \gamma^*$, the event

$$\|v_n + u_n\| < \gamma, n = n_0, n_0 + 1, \ldots, n_0 + K - 1 \tag{22}$$

is covered by the event

$$sign(v_{k_i}) \ne sign(u_{k_i}) \qquad i - 0,1,\cdots,K/N \tag{23}$$

Hence

$$\begin{aligned} &P\{\|v_n + u_n\| < \gamma \quad n = n_0, n_0 + 1, \cdots, n_0 + K - 1\} \\ &\le P\{sign(v_{k_i}) \ne sign(u_{k_i}) \quad i = 0,1,\cdots,K/N\} < \varepsilon \end{aligned} \tag{24}$$

now suppose

$$p_1 = P\{v_{k_i} > 0\} \quad p_2 = P\{u_{k_i} > 0\} \tag{25}$$

because v_n and u_n are independent for all n

$$\begin{aligned} &P\{sign(v_{k_i}) \ne sign(u_{k_i}) \quad i = 0,1,\cdots,K/N\} = \\ &\quad (p_1(1 - p_2) + p_2(1 - p_1))^{K/N} \end{aligned} \tag{26}$$

Hence under natural conditions that

$$0 < p_1 < 1 \ , \ \ 0 < p_2 < 1$$

it is obvious that for the given □ there exists M_1 such that

$$\left(p_1(1-p_2)+p_2(1-p_1)\right)^{M_1} < \varepsilon \qquad (27)$$

Choosing $M = M_1 N$ completes the proof.

REFERENCES

1. Fogel, E. and Huang, Y. F., "On the Value of Information in System Identification-Bounded Noise Case", ***Automatica***, Vol. 20, (1982), 229-238.
2. Goodwin, G. C., Sin, K. S., "Adaptive Filtering, Prediction and Control", Englewood Cliffs, NJ, Prentice-Hall, (1982).
3. Haykin, V., "Adaptive Filter Theory", 3rd Ed., Englewood Cliffs, NJ, Prentice-Hall, (1996).
4. Farhang-Borojeny, V., "Adaptive Filters: Theory and Applications", John-Wiley and Sons, U.K., (1998).
5. Dasgupta, S. and Huang, Y. F., "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, (May 1987), 383-392.
6. Hassibi, B., Sayed, A. H. and Kailath, T., "H^{∞} Optimality of the LMS Algorithm", ***IEEE Trans. on Signal Processing,*** Vol. 44, No. 2, (Feb. 1996), 267-280.
7. Rao, A. K. and Huang, Y. F., "Tracking Characteristics of an OBE Parameter Estimation Algorithm", ***IEEE Trans. on Signal Processing,*** Vol. 41, No. 3, (March 1993), 1140-1148.
8. Deller, J. R., Nayeri, M. and Odeh, S. F., "Least-Square Identification with Error Bounds for Real-Time Signal Processing and Control", ***Proceedings of the IEEE,*** Vol. 81, No. 6, (June 1993), 815-849.
9. Lin, V., Nayeri, V and Deller, J. R., "Automatic Bound Estimation: A Practical Development in Optimal Bounding Ellipsoid Processing", ***IEEE Signal Processing Letters,*** Vol. 4, No. 8, (August 1997), 236-239.
10. Gollamudi, S, Nagaraj, S., Kapoor, S. and Huang, Y. F., "Set-Membership Filtering and a Set-Membership Normalized LMS Algorithm with an Adaptive Step Size", ***IEEE Signal Processing Letters***, Vol. 5, No. 5, (May 1998), 111-114.
11. Durret, R., "Probability: Theory and Examples", Wadsworth and Brooks/Cole Publishing Co., California, (1993).
12. Doost-Hoseini, A. M. and Shahtalebi, K., "NLMS Algorithm with Variable Step-Size Using Set-Membership Identification", Submitted to ***Scientia Iranica***.
13. Lin, T. M., "Optimal Bounding Ellipsoid Algorithms with Automatic Bound Estimation", Ph.D Dissertation, Michigan State University, (1996).