A Set-Membership NLMS Algorithm with Time-Varying Error Bound

Juraci F. Galdino
IME – SE/3
Pç. Gen. Tibúrcio 80, Rio de Janeiro, RJ
BRAZIL, 22.290-270
Email: galdino@ime.eb.br

José A. Apolinário Jr.
IME – SE/3
Pç. Gen. Tibúrcio 80, Rio de Janeiro, RJ
BRAZIL, 22.290-270
Email: apolin@ieee.org

Marcello L. R. de Campos
COPPE/UFRJ – PEE
P.O.Box 68504, Rio de Janeiro, RJ
BRAZIL, 21.945-970
Email: campos@lps.ufrj.br

***Abstract*— Set Membership adaptive filtering is known for a number of attractive features, including reduction of computational complexity due to less frequent coefficient updates. This paper addresses the problem of choosing the error bound to be used in SM adaptation algorithms. This choice has often been based on the experience of the designer, and affects directly not only algorithm performance, but also its computational complexity. We propose a new time-varying error bound for a set-membership normalized least mean squares (SM-NLMS) algorithm that yields near constant average coefficient updating rate during both transient and steady state. The expressions given herein not only offer a new method to calculate the error bound automatically, but also describes the behavior of a conventional SM-NLMS algorithm. The results were obtained for the particular application of linear time-invariant channel estimation, but certainly provide insightful hints to other scenarios.**

I. Introduction

Set-membership adaptive filters (SM-AF) are known to be able to provide an alternative to conventional adaptive filtering with possible lower computational complexity and misadjustment, as well as faster convergence. In order to achieve this optimal performance, SM filtering must rely on a judicious choice of the error bound, γ, which is often application dependent. This has been a major difficulty to implement SM adaptive filters in some practical applications. In addition, critics of SM-AF have too often emphasized that computational complexity of SM-AF algorithms is hard to predict and control. These two factors have somewhat hindered the acceptance and clouded the impact that SM adaptive filters might have enjoyed and deserved. After all, set membership may yield adaptive filters the best of all worlds: better performance with lower complexity.

In this paper we investigate the application of a SM Normalized Least Mean Squares (SM-NLMS) algorithm for a time-invariant linear-channel estimation. For this particular application we present an analysis of the SM algorithm that resulted in an accurate expression for the probability of updating the coefficient vector. This analysis can be very useful in hardware design, for it can very accurately predict the overall computational power required per block of symbols in a training sequence. In addition, we proposed a new error bound that is proportional to the estimated mean squared error. The resulting algorithm has near constant probability of coefficient updates. Therefore we hope to have addressed two major obstacles to the implementation of SM-AF and also to have eased some of the concerns critics might have to these powerful adaptation methods.

II. NLMS Performance Analysis for Channel Estimation

In the scenario of interest in this paper, adaptive filters are employed to estimate the coefficients of time-invariant linear channels, which can also be viewed as a typical system-identification application. The noisy observation, at time instant k, is given by $d(k) = \mathbf{w}_o^T \mathbf{u}(k) + n(k)$, where $n(k)$ represents a zero-mean white Gaussian observation noise with variance σ_n^2, $\mathbf{w}_o$ represents the length-L linear-channel impulse-response vector $\mathbf{w}_o = [w_0 \ w_1 \ \cdots \ w_{L-1}]^T$, and $\mathbf{u}(k)$ represents the input-signal vector $\mathbf{u}(k) = [u(k) \ u(k-1) \ \cdots \ u(k-L+1)]^T$.

The input sequence $u(k)$ is assumed to consist of independent and equiprobable symbols ± 1 such that $\sigma_u^2 = 1$ and $\mathbf{u}^T(k)\mathbf{u}(k) = L$. Nevertheless, the results presented herein can be easily extended to a more general case of an M-ary phase-shift keying (M-PSK) constellation.

The NLMS estimate of $\mathbf{w}_o$, available at time k, is denoted by $\mathbf{w}(k)$ and is updated as follows:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{u}(k) \tag{1}$$

where μ is the step size parameter and $e(k)$ is the *a priori* estimation error at time k, given by $e(k) = d(k) - \mathbf{w}^T(k)\mathbf{u}(k)$. Notice that, for this particular application, LMS and NLMS algorithms are equivalent because the norm of the input signal vector $\mathbf{u}(k)$ is constant and known.

The coefficient-error vector and its mean squared norm are defined as $\Delta\mathbf{w}(k) = \mathbf{w}_o - \mathbf{w}(k)$ and $D(k) = \mathrm{E}[\Delta\mathbf{w}^T(k)\Delta\mathbf{w}(k)]$, respectively, where $\mathrm{E}[\cdot]$ denotes the expectation operator.

From the results presented in [1], assuming a time-invariant channel and using the *independence assumption* (filter coefficients are independent of the input signal), we have

$$D(k) = (1 - 2\mu + \mu^2 L)D(k-1) + (\mu^2\sigma_n^2 L) \tag{2}$$

Defining $\alpha = 1 - 2\mu + \mu^2 L$ and $\beta = \mu^2\sigma_n^2 L$, the expression for $D(k)$ can be rewritten as

$$D(k) = \alpha^k D(0) + \beta \sum_{i=0}^{k-1} \alpha^i \quad \text{where } \mathrm{D}(0) = ||\mathbf{w}_\mathrm{o}||^2$$

For $k \to \infty$ and $|\alpha| < 1$, we can write

$$D(\infty) = \frac{\mu L}{2 - \mu L}\ \sigma_n^2 \tag{3}$$

The mean-square output error (MSE) is given by

$$\xi(k) = D(k) + \sigma_n^2 = D(k) + \xi_{min} \tag{4}$$

for, in this case, the input-signal correlation matrix is the identity matrix.

According to the definitions presented in [2], the excess in the MSE is numerically equal to $D(k)$ and Eq. (3) gives the excess mean-square error. Equivalently, the misadjustment is given by $M = \frac{\mu L}{2-\mu L}$. Due to the fact that $|\alpha| < 1$, the step-size shall be in the range $0 < \mu < 2/L$. Given that $\xi(\infty) = (1+M)\xi_{min}$ then the step-size can be written as a function of the misadjustment:

$$\mu = \frac{2M}{(1+M)L} \tag{5}$$

III. A Set-Membership NLMS Algorithm

The formulation of the SM-NLMS algorithm [3] states that $\mathbf{w}(k)$ shall be updated only if the absolute value of the estimation error, $|e(k)|$, is greater than a pre-defined error bound, γ. The updated vector, $\mathbf{w}(k+1)$, should lead to an *a posteriori* error inside the boundary hiperplanes. The SM-NLMS algorithm chooses the closest boundary at a minimum distance from $\mathbf{w}(k)$. In this work, considering the application described in the previous section, we define the following updating scheme:

$$\mathbf{w}(k+1) = \begin{cases} \mathbf{w}(k) + \mu e(k)\mathbf{u}(k) & \text{if } |e(k)| > \gamma \\ \mathbf{w}(k) & \text{otherwise} \end{cases} \tag{6}$$

For the case $|e(k)| > \gamma$ in the equation above, the *a posteriori* error will lie inside the range $\pm\gamma$ as long as μ is chosen close to $1/L$; actually whenever $|1 - \mu L||e(k)| < \gamma$. Assuming that the *a priori* error $e(k)$ is a zero-mean Gaussian process with variance given by Eq.(4), the range of values of μ that guarantees probability p of having the new coefficient vector in the constraint set is given by

$$\mu \in [(1/L)(1-\rho),\ (1/L)(1+\rho)] \tag{7}$$

where $\rho = \frac{\gamma}{\sqrt{2\xi(k)}\text{erfc}^{-1}(1-p)}$ and $\text{erfc}[\cdot]$ denotes the *complementary error function* [4]. Notice that if $p = 1$, $\rho = 0$ and $\mu = 1/L$ (conventional NLMS update). When p is reduced, the range of possible values of μ increases. The updating scheme in Eq. (6), although a data-dependent filtering scheme, cannot strictly be considered an SM algorithm.

Assuming that the learning curve of the updating scheme proposed above can be approximated by the one generated from the fixed step-size algorithm of (1), $e(k)$ is modeled as a Gaussian random variable with zero mean and variance given by Eq. (4). As it will be seen at the end of this Section, the assumption mentioned above has proved valid for values of parameter γ commonly used in practical situations. The probability of updating at time instant k is given by

$$P[|e(k)| > \gamma] = \text{erfc}\left[\frac{\gamma}{\sqrt{2\xi(k)}}\right]$$

Bounds for the probability of update have already been used in the set-membership literature in order to compute the excess MSE after convergence [5].

We can assume that $e(l)$ and $e(m)$ are statistically independent for $l \neq m$ and define $z(k)$ as

$$z(k) = \begin{cases} 1 & \text{if } |e(k)| > \gamma \\ 0 & \text{otherwise} \end{cases}$$

and the instantaneous updating rate $r(k) = E[z(k)]$, which gives

$$r(k) = \text{erfc}\left[\frac{\gamma}{\sqrt{2\xi(k)}}\right] \tag{8}$$

for the particular application at hand.

For a block of B symbols, we can also define the block updating coefficient rate as

$$r_B = \frac{1}{B}\sum_{i=1}^{B} \text{erfc}\left[\frac{\gamma}{\sqrt{2\xi(i)}}\right]$$

It can be observed that the rate for coefficient update will be larger in the beginning of the adaptation and will become smaller as convergence is approached.

A. On the behavior of the error bound.

The analytical expression given in Eq. (8) for the updating rate may help the designer to obtain a practical value for γ. We remind the reader that choosing the upper bound for the estimation error has been a trial-and-error procedure and, as a rule of thumb, a value of $\gamma = \sqrt{5}\sigma_n$ has been often used.

For a particular misadjustment M, we can calculate μ according to Eq. (5), $\xi(\infty)$ from Eq. (4), and γ can be obtained from the updating rate after convergence, $r(\infty)$, using Eq. (8) as follows:

$$\gamma = erfc^{-1}[r(\infty)]\sqrt{2(M+1)}\sigma_n \tag{9}$$

The choice of γ, as seen in Eq. (9), is related to the desired rate of updating after convergence, the desired misadjustment, and the observation noise. However, in most practical situations the designer has only limited control over observation noise. Figure 1 shows $r(\infty) \times M$ for constant values of $\eta = \gamma^2/\sigma_n^2$. Values of $r(\infty)$ from 5% to 10% and M from 0.5 to 1 yield good speed of convergence and a value of η between 4 and 5, typically found in the set-membership filtering literature, i.e., $\sqrt{4}\sigma_n \leq \gamma \leq \sqrt{5}\sigma_n$.

B. Verifying the applicability of the NLMS analysis in the SM framework.

In the previous subsection, we have assumed that the main expression of the analysis of the conventional NLMS, i.e., the mean squared norm of the coefficient error vector expressed in Eq. (2), was valid for the SM-NLMS algorithm. In order to check the validity of this assumption, we ran an experiment consisting of a channel identification with an impulse response given by [4] $\mathbf{h} = [0.2270\ 0.4601\ 0.6881\ 0.4601\ 0.2270]^T$ and the training set being a block of 500 independent and equiprobable symbols $u(k) = \pm 1$.

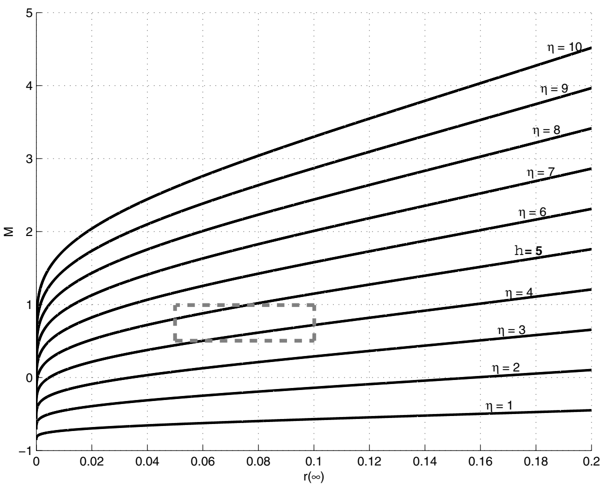

Fig. 1. Curves of $r(\infty) \times M$ for fixed values of η.

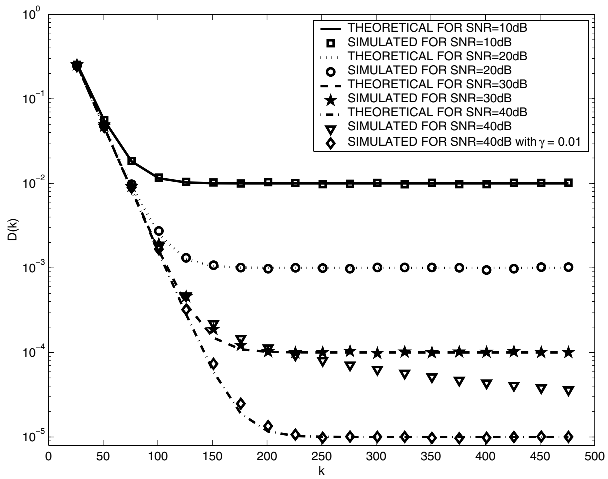

Fig. 2. Theoretical and simulated $D(k)$ for different values of SNR.

Figure 2 shows the behavior of $D(k)$ averaged over 1000 independent runs for values of SNR from 10dB to 40dB. Except for the case of $SNR = 40$dB, simulations and theoretical curves agree very well. In this experiment, we have set γ equal to 0.0316 and M equal to 0.1, such that $\mu = 0.0364$ for all values of SNR. According to the equations presented in Sections II and III, for $\gamma = 0.0316$ and the values of SNR equal to 10dB, 20dB, 30dB, and 40dB, $r(\infty)$ should be equal to 0.924, 0.763, 0.3404, and 0.0026, respectively. For $SNR = 40$dB the value of $\gamma = 0.0316$ was clearly inappropriate for it yielded a very low rate of updating ($r(\infty) = 0.26\%$), which was the reason for the only unmatched curve. For the same scenario with $SNR = 40$dB but with $\gamma = 0.01$, the updating rate was 48%, leading to a very good match between theoretical and simulated curves in Figure 2.

In a second experiment, the same setup was used in order to verify the accuracy of the estimates of $r(\infty)$ for values of SNR from 5dB to 40dB and for different values of γ. We measured the rate of coefficient adaptation after convergence, $r(\infty)$, for three different values of γ. The results are presented

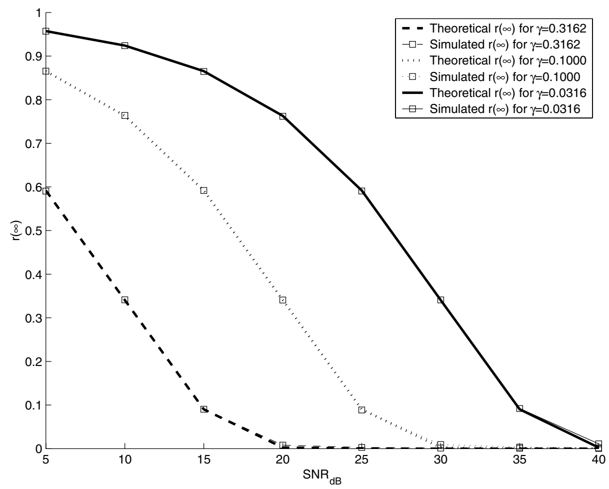

Fig. 3. Theoretical and simulated values for $r(\infty)$ as a function of the SNR.

in Figure 3. Simulation results and theoretical values agreed very well.

It is worth mentioning that other experiments, not shown in this paper due to lack of space, were carried out and pointed out that, for the usual range of η, the analytical expressions obtained from the analysis of the NLMS algorithm, for this particular application, accurately describe the behavior of the SM-NLMS algorithm presented in Eq.(6).

IV. A NOVEL TIME-VARYING ERROR BOUND WITH NEAR CONSTANT UPDATING RATE

SM algorithms usually have high updating rate during convergence and low updating rate after convergence. It may be that an approximately constant updating rate is desirable; in that case, an alternative approach to the choice of γ is necessary. Our proposition is to make the error bound time-varying, $\gamma(k)$, and proportional to $\xi(k)$. In the beginning of the adaptation process, the larger threshold will cause less frequent coefficient updates. As the adaptive filter converges and the average estimation error at each time instant becomes smaller, an also smaller threshold will render less misadjustment.

For $\gamma(k) = \sqrt{\tau\xi(k)}$ with $\xi(k)$ given by Eq. (4), the probability of updating the coefficients at time instant k is equal to

$$P[|e(k)| > \gamma(k)] = r(k) = \text{erfc}\left[\sqrt{\frac{\tau}{2}}\right]$$

The constant probability of updating may be advantageous as one may expect the same average computational complexity throughout the training set.

Table I shows the equations of the proposed algorithm. Values of $\gamma(k)$ can be calculated in advance and stored in memory during algorithm initialization.

For this algorithm, although the value of μ is constant, the chosen bounds are time-varying and data-dependent. The corresponding ρ (as in (7)) becomes

$$\rho = \frac{\sqrt{\tau}}{\text{erfc}^{-1}(1-p)}$$

TABLE I

THE SM-NLMS ALGORITHM WITH CONSTANT UPDATING RATE.

SM-NLMS CU

Initialization

Choose r (the desired rate of updating):

$\tau = 2[erfc^{-1}(r)]^2;$

Choose M (the desired misadjustment):

$\mu = \frac{2M}{(1+M)L}; \quad \alpha = (1 - 2\mu + \mu^2 L); \quad \beta = 2\mu\sigma^2;$

Running the algorithm

for each k

{

$e(k) = d(k) - \mathbf{w}^T(k)\mathbf{u}(k);$

$\xi(k) = \alpha\xi(k-1) + \beta;$

$\gamma(k) = \sqrt{\tau\xi(k)};$

if $|e(k)| > \gamma(k)$

{

$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{u}(k);$

}

else

{

$\mathbf{w}(k+1) = \mathbf{w}(k);$

}

}

which allows a simple calculation of the probability of having the *a posteriori* error lying inside the constraint set. In other words, p gives the probability that the algorithm updates as a "true" SM algorithm.

V. SIMULATION RESULTS

In this section, we present the results of an experiment carried out in order to show the performance of the proposed algorithm in the same scenario previously described. We obtained $\mu = 0.1333$ for a misadjustment $M = 0.5$. We ran the same channel identification experiment for few different updating rates, from 5% to 50%, with a constant SNR equal to 30dB. Figure 4 shows the simulation results averaged over 5000 independent runs. For $r(\infty) = 50\%$, $p = 94.6\%$, i.e., given that an update was required, 94.6% of the updates rendered a solution inside the constraint set. For smaller values of $r(\infty)$, the probability of updating to $\mathbf{w}(k+1)$ inside the constraint set rapidly approached 100%. In Figure 4 the updating rate is not exactly constant during initial convergence. The difference is particularly noticeable when the desired rate is small. This is due to the fact that, during convergence and for small updating rates, the analysis expressions are only rough approximations of the real curves. This can be seen in Figure 5: the curves associated with smaller values of $r(k)$ are further apart from the theoretical curve.

VI. CONCLUSIONS

In this paper we evaluated the rate of coefficient updates for an SM-NLMS algorithm when applied in a time-invariant linear-channel estimation. An accurate analysis for the case of fixed error bound was provided based in a previous work and assuming a typical communications application. We also proposed a new approach to the choice of the error bound to be used by the SM algorithm where an estimate of the mean squared error is used. This choice renders a probability of coefficient update which is approximately constant and, therefore, the algorithm does not "suffer" from the unbalanced computational load over time that is seen in conventional SM adaptive filters.

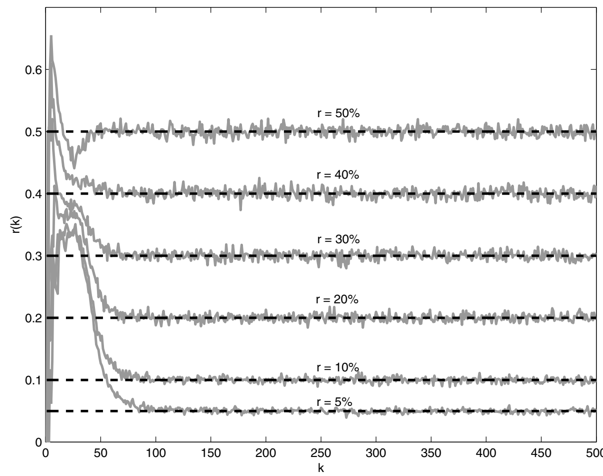

Fig. 4. Curves of updating rates with different values of τ.

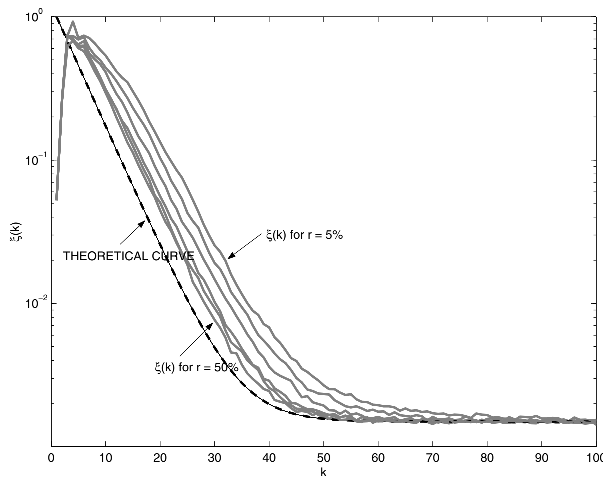

Fig. 5. Curves of MSE with different values of r.

REFERENCES

[1] J. F. Galdino, E. L. Pinto, and M. S. de Alencar, "Analytical performance of the LMS algorithm on the estimation of wide sense stationary channels, "*IEEE Transactions on Communications*, vol. 52, no. 6, pp. 982–991, June 2004.

[2] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementation*, 2nd ed. Norwell, Massachusetts, USA: Kluwer Academic Publishers, 2002.

[3] S. Gollamudi, S. Nagaraj, S. Kapoor, and Y.-F. Huang, "Set-Membership filtering and a set-membership normalized LMS algorithm with an adaptive step size," *IEEE Signal Processing Letters*, vol. 5, no. 5, pp. 111–114, May 1998.

[4] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[5] P. S. R. Diniz and Stefan Werner, "Set-Membership Binormalized Data-Reusing LMS Algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 124–134, January 2003.

ACKNOWLEDGMENT

The authors would like to thank Brazilian agencies CNPq, FAPERJ, and CAPES for partial funding of this work, and the anonymous reviewers for their valuable comments.