CONVERGENCE ANALYSIS OF THE $\varepsilon$ NSRLMMN ALGORITHM

Mohammed Mujahid Ulla Faiz and Azzedine Zerguine

Department of Electrical Engineering
King Fahd University of Petroleum & Minerals
Dhahran 31261, Saudi Arabia
E-mail: {mujahid, azzedine}@kfupm.edu.sa

ABSTRACT

In this work, the $\varepsilon$–normalized sign regressor least mean mixed-norm (NSRLMMN) adaptive algorithm is proposed. The proposed algorithm exhibits increased convergence rate as compared to the least mean mixed-norm (LMMN) and the sign regressor least mean mixed-norm (SRLMMN) algorithms. Also, the steady-state analysis and convergence analysis are presented. Moreover, the proposed $\varepsilon$–NSRLMMN algorithm substantially reduces the computational load, a major drawback of the $\varepsilon$–normalized least mean mixed-norm (NLMMN) algorithm. Finally, simulation results are presented to support the theoretical findings.

Keywords: Adaptive filters, LMS, LMF, Least Mean Mixed-Norm (LMMN), Sign regressor LMMN algorithm.

1. INTRODUCTION

While the least mean mixed-norm (LMMN) algorithm was introduced in order to combine the advantages of both the least mean square (LMS) and the least mean fourth (LMF) algorithms [1]–[6], the sign adaptive filters were proposed in order to reduce the computational cost and to simplify the hardware implementation [7]–[8]. However, these sign adaptive filters result in slower convergence speeds due to clipping of the estimation error or the input data, or both [9]. The algorithm based on clipping of the input data of the LMMN is known as the sign regressor least mean mixed-norm (SRLMMN) algorithm. The convergence speed of the SRLMMN algorithm can be increased by normalizing it. Hence the name the $\varepsilon$–normalized sign regressor least mean mixed-norm (NSRLMMN) algorithm. From the simulation results it is shown that the $\varepsilon$–NSRLMMN algorithm outperforms both the LMMN and SRLMMN algorithms.

The paper is organized as follows. In Section 2, the $\varepsilon$–NSRLMMN algorithm is proposed. The steady-state analysis of the proposed algorithm is derived in Section 3, and Section 4 presents its convergence analysis. A comparison of the computational complexity of the proposed algorithm with some of other algorithms in the family is presented in Section 5. Finally, the simulation results and conclusions are presented in Sections 6 and 7, respectively.

2. THE $\varepsilon$–NSRLMMN ALGORITHM

Consider a zero-mean random variable $d$ with realizations $\{d(0), d(1), \ldots\}$, and a zero-mean random row vector $u$ with realizations $\{u_0, u_1, \ldots\}$. The LMMN algorithm is based on the following convex cost function [1]–[3]:

$$J_i = E\left[\delta e_i^2 + (1 - \delta) e_i^2\right], \quad 0 \leq \delta \leq 1. \quad (1)$$

where $\delta$ is the mixing parameter and $e_i$ denotes the estimation error given by

$$e_i = d_i - u_i w_{i-1}. \quad (2)$$

The update equation for the $\varepsilon$–NSRLMMN algorithm can be shown to be governed by the following recursion:

$$w_i = w_{i-1} + \frac{\mu}{\varepsilon + |u_i|} \text{sign}[u_i^T e_i (\delta + (1 - \delta) e_i^2)], \quad i \geq 0, \quad \mu \geq 0. \quad (3)$$

where $w_i$ (column vector) is the updated weight vector at time $i$ with optimal weight vector $w^o$, $\mu$ is the step-size, $\varepsilon$ is a small positive constant used for regularization purposes, $|u_i| > 0$, and $H[u_i]$ is some positive-definite Hermitian matrix-valued function of $u_i$ defined by

$$H[u_i] = \text{diag}\left\{\frac{1}{|u_i|}, \frac{1}{|u_i|}, \ldots, \frac{1}{|u_i|}\right\}. \quad (4)$$

where $M$ is the filter length and $\text{sign}[u_i] = H[u_i] u_i^T$.

3. STEADY-STATE ANALYSIS OF THE $\varepsilon$–NSRLMMN ALGORITHM

We shall assume that the data $\{d_i, u_i\}$ satisfy the following assumptions of the stationary data model [10]:

A.1 There exists an optimal weight vector $w^o$ such that $d_i = u_i w^o + v_i$.

A.2 The noise sequence $v_i$ is independent and identically distributed (i.i.d.) with variance $\sigma_v^2 = E[v_i^2]$ and is independent of $u_j$ for all $i, j$.

A.3 The initial condition $w_{-1}$ is independent of the zero mean random variables $\{d_i, u_i, v_i\}$.

A.4 The regressor covariance matrix is $R = E[u_i^T u_i] > 0$.

For the adaptive filter of the form in (3), and for any data $\{d_i, u_i\}$, assuming filter operation in steady-state, the following variance relation holds [10]:

$$\mu E\left[|u_i|^2; \sigma_v^2; e_i\right] = 2E[e_o g(e_o)] \quad \text{as } i \to \infty, \quad (5)$$

where

$$E[|u_i|^2] = E[u_i H[u_i] u_i^T], \quad (6)$$

$$e_i = e_o + v_i, \quad (7)$$

with $g[e]$ denoting some function of $e$, and $e_o = u_i (w^o - w_{i-1})$ is the a priori estimation error. Then $g[e_i]$ for the
\[ g[\xi] = e_\xi \left[ (1 - 3) \xi^2 \right] \]

\[ = \delta (e_\xi + v_1) \left[ e_\xi^4 + e_\xi v_1^2 + 2 e_\xi^3 v_1 \right] \]

\[ + \delta \left[ e_\xi^4 + e_\xi^3 v_1 + e_\xi^3 v_1^2 + 15 e_\xi^2 v_1^3 + 20 e_\xi v_1^4 + v_1^5 \right] \]

\[ + \frac{2 \delta^2}{(\varepsilon + ||u||_H^2)} \left[ e_\xi^4 + e_\xi^3 v_1 + 4 e_\xi v_1^3 + v_1^4 \right] \]

\[ + 4 e_\xi v_1^3 + v_1^4 \]

\[ \text{(8)} \]

where \( \delta = 1 - \delta \). By using the fact that \( e_\xi \) and \( v_1 \) are independent, we reach at the following expression for the term \( E[e_\xi g[\xi]] \):

\[ E[e_\xi g[\xi]] = \delta E \left[ \frac{e_\xi^4}{\varepsilon + ||u||_H^2} \right] + (\delta + 3 \delta \sigma_\varepsilon^2) \]

\[ \times E \left[ \frac{e_\xi^2}{\varepsilon + ||u||_H^2} \right]. \]

\[ \text{(9)} \]

Ignoring third and higher-order terms of \( e_\xi \), we obtain

\[ E[e_\xi g[\xi]] \approx (\delta + 3 \delta \sigma_\varepsilon^2) E \left[ \frac{e_\xi^2}{\varepsilon + ||u||_H^2} \right]. \]

\[ \text{(10)} \]

To evaluate the term \( E[||u||_H^2 g[\xi]] \), we start by noting that

\[ g^2[\xi] = \frac{\delta^2}{(\varepsilon + ||u||_H^2)} \left[ e_\xi^2 + v_1^2 + 2 e_\xi v_1 \right] \]

\[ = \frac{\delta^2}{(\varepsilon + ||u||_H^2)} e_\xi^2 + 6 e_\xi v_1 + 6 e_\xi v_1^2 \]

\[ + 15 e_\xi^2 v_1^3 + 20 e_\xi v_1^4 + v_1^5 \]

\[ + \left( \frac{2 \delta^2}{(\varepsilon + ||u||_H^2)} \right) e_\xi^2 + 6 e_\xi v_1^2 + 4 e_\xi^3 v_1 \]

\[ + 4 e_\xi v_1^3 + v_1^4. \]

\[ \text{(11)} \]

In order to simplify (13), we use the separation principle, namely, that at steady-state, \( ||u||_H^2 \) is independent of \( e^2_\xi \). Therefore, we obtain

\[ \mu(\delta^2 + 15 \delta^2 \xi^4 + 12 \delta \delta \sigma_\varepsilon^2) E \left[ \frac{||u||_H^2}{(\varepsilon + ||u||_H^2)^2} \right] E[e_\xi^2] \]

\[ + \mu(\delta^2 \sigma_\varepsilon^2 + \bar{\delta} \xi^4 + 2 \delta \delta \xi^5) E \left[ \frac{||u||_H^2}{(\varepsilon + ||u||_H^2)^2} \right] \]

\[ = 2(\delta + 3 \delta \sigma_\varepsilon^2) E \left[ \frac{e_\xi^2}{\varepsilon + ||u||_H^2} \right]. \]

\[ \text{(14)} \]

which can be set up compactly as

\[ \mu(\delta^2 \sigma_\varepsilon^2 + \bar{\delta} \xi^4 + 2 \delta \delta \xi^5) \mathcal{Z}_1 = 2(\delta + 3 \delta \sigma_\varepsilon^2) \mathcal{Z}_2 \]

\[ - \mu(\delta^2 + 15 \delta^2 \xi^4 + 12 \delta \delta \sigma_\varepsilon^2) \mathcal{Z}_2 \]

\[ \text{where} \]

\[ \mathcal{Z}_1 = \frac{1}{(\varepsilon + ||u||_H^2)^2}, \]

\[ \mathcal{Z}_2 = \frac{1}{(\varepsilon + ||u||_H^2)^2}. \]

\[ \text{(15)} \]

Therefore, the expression for the steady-state excess-mean-square error (EMSE) \( \zeta = E[e^2_\xi] \) of the \( \varepsilon \)–NSRLMMN algorithm is given by

\[ \zeta = \frac{\mu(\delta^2 \sigma_\varepsilon^2 + \bar{\delta} \xi^4 + 2 \delta \delta \xi^5)}{2(\delta + 3 \delta \sigma_\varepsilon^2) \mathcal{Z}_2 - \mu(\delta^2 + 15 \delta^2 \xi^4 + 12 \delta \delta \sigma_\varepsilon^2) \mathcal{Z}_1}. \]

\[ \mathcal{Z}_1 \triangleq \mathcal{Z}_2 = E \left[ \frac{1}{||u||_H^2} \right]. \]

\[ \text{(16)} \]

(18) becomes

\[ \zeta = \frac{\mu(\delta^2 \sigma_\varepsilon^2 + \bar{\delta} \xi^4 + 2 \delta \delta \xi^5)}{2(\delta + 3 \delta \sigma_\varepsilon^2) - \mu(\delta^2 + 15 \delta^2 \xi^4 + 12 \delta \delta \sigma_\varepsilon^2)}, \]

which is independent of the regressor.

An alternative expression for the steady-state EMSE of the \( \varepsilon \)–NSRLMMN algorithm can be obtained by using the assumption \( \varepsilon \approx 0 \) in order to simplify (13) into

\[ \mu(\delta^2 + 15 \delta^2 \xi^4 + 12 \delta \delta \sigma_\varepsilon^2) E \left[ \frac{e_\xi^2}{(\varepsilon + ||u||_H^2)^2} \right] + \mu(\delta^2 \sigma_\varepsilon^2 + \bar{\delta} \xi^4 + 2 \delta \delta \xi^5) \]

\[ + 2 \delta \delta \xi^4 E \left[ \frac{1}{||u||_H^2} \right] = 2(\delta + 3 \delta \sigma_\varepsilon^2) E \left[ \frac{e_\xi^2}{(\varepsilon + ||u||_H^2)^2} \right]. \]

\[ \text{(21)} \]

Now, let us use the following steady-state approximation:

\[ E \left[ \frac{e_\xi^2}{||u||_H^2} \right] \approx E \left[ \frac{e_\xi^2}{||u||_H^2} \right]. \]

\[ \text{(22)} \]

In [11], we have shown that

\[ E[||u||_H^2] = \sqrt{\frac{2}{\pi \sigma_\varepsilon^2} \text{Tr}(R)}. \]

\[ \text{(23)} \]
Substituting (22) and (23) into (21) we get
\[\mu(\delta^2\sigma_1^2 + \delta^2\xi_1^2 + 2\delta\xi_1^2)\frac{1}{\|u_i\|^2} = \frac{2(\delta + 3\delta\sigma_1^2)}{\mu}. \tag{24}\]
Therefore, the steady-state EMSE of the \(\varepsilon\)-NSRLMMN algorithm can also be approximated by
\[\zeta = \mu \text{Tr}((\delta^2\sigma_1^2 + \delta^2\xi_1^2 + 2\delta\xi_1^2)\frac{1}{\|u_i\|^2})\frac{1}{2\sigma_u^2} \sqrt{\frac{\text{E}[\xi_u]}{2\text{Tr}(R)}}. \tag{25}\]
Ultimately, an expression for the mean-square error (MSE) of the \(\varepsilon\)-NSRLMMN algorithm is given by
\[\text{E}[^2] = \zeta + \sigma_e^2. \tag{26}\]

4. CONVERGENCE ANALYSIS OF THE \(\varepsilon\)-NSRLMMN ALGORITHM

In [1], the approximate bound on the step-size of the LMMN algorithm was obtained by simply combining the step-size bounds of LMS and LMF. Similarly, the approximate bound on the step-size of our proposed \(\varepsilon\)-NSRLMMN algorithm can be obtained by combining the step-size bounds of \(\varepsilon\)-NSRLMS and \(\varepsilon\)-NSRLMF.

It was shown in [12] that the convergence in the mean for the \(\varepsilon\)-NSRLMS algorithm is guaranteed by the stability condition for the \(\varepsilon\)-LMM algorithm, namely,
\[0 < \mu_{\varepsilon\text{-NSRLMS}} < 2. \tag{27}\]
Also, the mean convergence of the \(\varepsilon\)-NSRLMF algorithm can be bounded by
\[0 < \mu_{\varepsilon\text{-NSRLMF}} < \mu_{\text{upper}}. \tag{28}\]
Thus, by combining (27) and (28) the mean convergence of the \(\varepsilon\)-NSRLMMN algorithm can be approximated by
\[0 < \mu_{\varepsilon\text{-NSRLMMN}} < 2\delta + (1 - \delta)\mu_{\text{upper}}. \tag{29}\]
From (29) it is clear that the \(\varepsilon\)-NSRLMMN algorithm reduces to \(\varepsilon\)-NSRLMF and \(\varepsilon\)-NSRLMS algorithms when \(\delta = 0\) and \(\delta = 1\), respectively. Our future work will focus on finding the upper bound for the step-size of the \(\varepsilon\)-NSRLMF algorithm.

5. COMPUTATIONAL COMPLEXITY

In this section, the computational complexity of the \(\varepsilon\)-NSRLMMN algorithm is compared with that of other algorithms in the family, e.g., LMMN, SRLMMN, and \(\varepsilon\)-NLMN algorithms. Tables 1 and 2 present this comparison for real- and complex-valued data, respectively, in terms of the number of real additions (+), real multiplications (×), real divisions (/), and comparisons with zero per iteration. Moreover, \(M\) is the filter order. As can be seen from Table 1, the real-valued data case, both the \(\varepsilon\)-NLMN and \(\varepsilon\)-NSRLMMN algorithms have similar computational complexity, while there are 2\(M\) and 2\(M\) + 4 extra additions and multiplications per iteration, respectively, for the \(\varepsilon\)-NLMN algorithm when compared to the \(\varepsilon\)-NSRLMMN algorithm in the complex-valued data case as reported from Table 2.

### 6. SIMULATION RESULTS

In order to evaluate the steady-state and convergence performance of our proposed algorithm, extensive simulations are carried out for this purpose. The parameter settings in this study are as follows. In all the simulations, we have chosen \(\varepsilon = 10^{-6}\), the mixing parameter is fixed at \(\delta = 0.5\) (except Figures 5-6), and the filter length is fixed at \(M = 10\) for Figures 1-2 and \(M = 5\) for Figures 3-8.

First, the steady-state MSE of the \(\varepsilon\)-NSRLMMN algorithm using white and correlated Gaussian regressors is shown in Figures 1-2, respectively. In Figure 2, the correlated data is obtained by passing a unit-variance i.i.d. Gaussian data through a first-order auto-regressive model with transfer function \(\sqrt{1 - \alpha^2}\frac{1}{\tau + s}\) and \(\alpha = 0.8\). In Figures 1-2, the MSE is plotted as a function of the step-size \(\mu\) in additive white Gaussian noise (AWGN) environment for a signal-to-noise ratio (SNR) of 30 dB. As observed from these figures, the simulation results are in a good match with the theoretical results ((20) and (25)), which are, respectively, the first and second approximations of the steady-state EMSE of the \(\varepsilon\)-NSRLMMN algorithm. Also, as can be seen from these figures, the theoretical results are in a better match with the simulation results for correlated Gaussian data than white Gaussian data.

Second, the convergence behavior of the \(\varepsilon\)-NSRLMMN algorithm is compared with that of LMMN, SRLMMN, and \(\varepsilon\)-NLMN algorithms in an unknown system identification setup with
\[w^* = [0.227 0.460 0.688 0.460 0.227]^T. \tag{30}\]
Figure 3 shows the convergence performance of all the four algorithms using white Gaussian regressors in a uniform noise environment with SNR = 10 dB. As is depicted from this figure, the \(\varepsilon\)-NSRLMMN algorithm results in superior performance over the LMMN and SRLMMN algorithms, but is only slightly inferior when compared to the \(\varepsilon\)-NLMN algorithm. Also, it is interesting to note that the performance of the SRLMMN algorithm is found to be identical to that of the LMMN algorithm for the same misadjustment. No deterioration has occurred to the SRLMMN algorithm. One can also observe this particular behavior from Figure 4, which shows the comparison of the third-tap weight learning curves of all the four algorithms for the same parameter settings.

### Table 1: Computational cost for real-valued data.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>+</th>
<th>×</th>
<th>/</th>
<th>sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMMN</td>
<td>2(M) + 2</td>
<td>2(M) + 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRLMMN</td>
<td>2(M) + 2</td>
<td>2(M) + 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\varepsilon)-NLMN</td>
<td>4(M) + 4</td>
<td>4(M) + 4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>(\varepsilon)-NSRLMMN</td>
<td>4(M) + 2</td>
<td>4(M) + 2</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Computational cost for complex-valued data.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>+</th>
<th>×</th>
<th>/</th>
<th>sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMMN</td>
<td>8(M) + 3</td>
<td>8(M) + 6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRLMMN</td>
<td>6(M) + 3</td>
<td>6(M) + 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\varepsilon)-NLMN</td>
<td>10(M) + 3</td>
<td>10(M) + 6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>(\varepsilon)-NSRLMMN</td>
<td>8(M) + 3</td>
<td>8(M) + 2</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
Third, Figures 5-6 demonstrate, respectively, the MSE and normalized weight error vector learning behaviors of the \( \varepsilon \)-NSRLMMN algorithm for different values of the mixing parameter \( \delta \) in a uniform noise environment at SNR = 10 dB. As can be seen from these figures, the \( \varepsilon \)-NSRLMMN algorithm boils down to \( \varepsilon \)-NSRLMF and \( \varepsilon \)-NSRLMS algorithms when \( \delta = 0 \) and \( \delta = 1 \), respectively. Therefore, by controlling \( \delta \) we can control the tradeoff between fast convergence rate and small misadjustment. We also find that for uniform noise, the \( \varepsilon \)-NSRLMF algorithm is superior to both \( \varepsilon \)-NSRLMS and \( \varepsilon \)-NSRLMMN algorithms.

Finally, Figures 7-8 illustrate, respectively, the MSE and normalized weight error vector convergence behaviors of the \( \varepsilon \)-NSRLMMN algorithm in uniform, Gaussian and Laplacian noise environments for SNR = 10 dB. As can be seen from Figure 7 that the best performance in terms of convergence behavior is obtained with uniform noise while the worst performance is obtained with Laplacian noise. We also note from Figure 8 that the lowest weight error is reached by the proposed algorithm for uniform noise environment as compared to Gaussian and Laplacian noise environments.

7. CONCLUSIONS

In this work, the \( \varepsilon \)-NSRLMMN algorithm is presented and resulted in a significant reduction in computational load over the \( \varepsilon \)-NLMMN algorithm. The proposed \( \varepsilon \)-NSRLMMN algorithm has been shown to exhibit slightly slower convergence rate than the \( \varepsilon \)-NLMMN algorithm for the same steady-state error. The mean-square analysis of the \( \varepsilon \)-NSRLMMN algorithm is performed and is found to corroborate the simulation results. Also, the convergence behavior of the proposed algorithm is analyzed for different values of the mixing parameter and different noise environments.

Acknowledgment: The authors would like to acknowledge the support provided by King Fahd University of Petroleum & Minerals.

REFERENCES


Figure 1: MSE performance of the \( \varepsilon \)-NSRLMMN algorithm using white Gaussian regressors in AWGN environment with SNR = 30 dB.

Figure 2: MSE performance of the \( \varepsilon \)-NSRLMMN algorithm using correlated Gaussian regressors in AWGN environment with SNR = 30 dB.
Figure 3: Comparison of the MSE learning curves of LMMN, SRLMMN, $\epsilon$–NLMMN, and $\epsilon$–NSRLMMN algorithms in a uniform noise environment with SNR = 10 dB.

Figure 4: Comparison of the third-tap weight learning curves of LMMN, SRLMMN, $\epsilon$–NLMMN, and $\epsilon$–NSRLMMN algorithms in a uniform noise environment with SNR = 10 dB.

Figure 5: Comparison of the MSE learning curves of the $\epsilon$–NSRLMMN algorithm for different values of $\delta$ in a uniform noise environment with SNR = 10 dB.

Figure 6: Comparison of the normalized weight error vector learning curves of the $\epsilon$–NSRLMMN algorithm for different values of $\delta$ in a uniform noise environment with SNR = 10 dB.

Figure 7: Comparison of the MSE learning curves of the $\epsilon$–NSRLMMN algorithm in uniform, Gaussian and Laplacian noise environments with SNR = 10 dB.

Figure 8: Comparison of the normalized weight error vector learning curves of the $\epsilon$–NSRLMMN algorithm in uniform, Gaussian and Laplacian noise environments with SNR = 10 dB.