Research Article

Analysis of the Sign Regressor Least Mean Fourth Adaptive Algorithm

Mohammed Mujahid Ulla Faiz, Azzedine Zerguine (EURASIP Member), and Abdelmalek Zidouri

Electrical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia

Correspondence should be addressed to Azzedine Zerguine, azzedine@kfupm.edu.sa

A novel algorithm, called the signed regressor least mean fourth (SRLMF) adaptive algorithm, that reduces the computational cost and complexity while maintaining good performance is presented. Expressions are derived for the steady-state excess-mean-square error (EMSE) of the SRLMF algorithm in a stationary environment. A sufficient condition for the convergence in the mean of the SRLMF algorithm is derived. Also, expressions are obtained for the tracking EMSE of the SRLMF algorithm in a nonstationary environment, and consequently an optimum value of the step-size is obtained. Moreover, the weighted variance relation has been extended in order to derive expressions for the mean-square error (MSE) and the mean-square deviation (MSD) of the proposed algorithm during the transient phase. Computer simulations are carried out to corroborate the theoretical findings. It is shown that there is a good match between the theoretical and simulated results. It is also shown that the SRLMF algorithm has no performance degradation when compared with the least mean fourth (LMF) algorithm. The results in this study emphasize the usefulness of this algorithm in applications requiring reduced implementation costs for which the LMF algorithm is too complex.

1. Introduction

Reduction in complexity of the least mean square (LMS) algorithm has always received attention in the area of adaptive filtering [1–3]. This reduction is usually done by clipping either the estimation error or the input data, or both to reduce the number of multiplications necessary at each algorithm iteration. The algorithm based on clipping of the estimation error is known as the sign error or more commonly the sign algorithm (SA) [4–8], the algorithm based on clipping of the input data is known as the sign regressor algorithm (SRA) [9–12], and the algorithm based on clipping of both the estimation error and the input data is known as the sign sign algorithm (SSA) [13, 14]. These algorithms result in a performance loss when compared with the conventional LMS algorithm [9, 10]. However, significant reduction in computational cost and simplified hardware implementation can justify this poor performance in applications requiring reduced implementation costs [15, 16].

The behavior of the SRA algorithm depends on the input data. It is shown in [11] that for some inputs, the LMS algorithm is stable while the SRA algorithm is unstable. This is a drawback of the SRA algorithm when compared with the SA algorithm since the latter is more stable than the LMS algorithm [4, 16]. The SRA algorithm is always stable when the input data is Gaussian as in the case of speech processing. Also, the performance of the SRA algorithm is superior to that of the SA algorithm for Gaussian input data. It is shown in [10] that the SRA algorithm is much faster than the SA algorithm in achieving the desired steady-state mean-square error for white Gaussian data. Theoretical studies of the SRA algorithm with correlated Gaussian data in both stationary and nonstationary environments are found in [12].

The convergence rate and the steady-state mean-square error of the SRA algorithm is only slightly inferior to those of the LMS algorithm for the same parameter setting. In [10], the convergence rate of the SRA algorithm is compared with that of the LMS algorithm to show that the SRA algorithm converges slower than the LMS algorithm

by a factor of $2/\pi$ for the same steady-state mean-square error.

It is shown in [17] that the SRA algorithm exhibits significantly higher robustness against the impulse noise than the LMS algorithm.

The above-mentioned advantages motivate us to analyze the proposed sign regressor least mean fourth (SRLMF) adaptive algorithm. In this paper, the mean-square analysis, the convergence analysis, the tracking analysis, and the transient analysis of the SRLMF algorithm are carried out. The framework used in this work relies on energy conservation arguments [18]. Expressions are evaluated for the steady-state excess-mean-square error (EMSE) of the SRLMF algorithm in a stationary environment. A condition for the convergence of the mean behavior of the SRLMF algorithm is also derived. Also, expressions for the tracking EMSE in a nonstationary environment are presented. An optimum value of the step-size μ is also evaluated. Moreover, an extension of the weighted variance relation is provided in order to derive expressions for the mean-square error (MSE) and the mean-square deviation (MSD) of the proposed algorithm during the transient phase. From the simulation results it is shown that both the SRLMF algorithm and the least mean fourth (LMF) algorithm [19] have a similar performance for the same steady-state EMSE. Moreover, the results show that the theoretical and simulated results are in good agreement.

The paper is organized as follows: following the Introduction is Section 2 where the proposed algorithm is developed, while the mean-square analysis of the proposed SRLMF algorithm is presented in Section 3. The convergence analysis of the proposed algorithm is presented in Section 4. Section 5 presents the tracking analysis of the proposed algorithm for random walk channels and as a by-product of this analysis the optimum value of step-size for these channels is derived. And Section 6 presents thoroughly the transient analysis of the proposed algorithm. The Computational Load is detailed in Section 7. To investigate the performance of the proposed algorithm, several simulation results for different scenarios are presented in Section 8. Finally, some conclusions are given in Section 9.

2. Algorithm Development

The SRLMF algorithm is based on clipping of the regression vector $\mathbf{u}_i$ (row vector). Consider now the adaptive filter, which updates its coefficients according to the following recursion [18]:

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathrm{H}[\mathbf{u}_i]\mathbf{u}_i^* \mathrm{g}[e_i], \quad i \geq 0, \tag{1}$$

where $\mathbf{w}_i$ (column vector) is the updated weight vector at time i, μ is the step-size, $\mathrm{H}[\mathbf{u}_i]$ is some positive-definite Hermitian matrix-valued function of $\mathbf{u}_i$, $\mathrm{g}[e_i]$ denotes some function of the estimation error signal given by

$$e_i = d_i - \mathbf{u}_i \mathbf{w}_{i-1}, \tag{2}$$

where d_i is the desired signal. When the data is real-valued and $\mathrm{g}[e_i] = e_i^3$, the general update form in (1) becomes

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathrm{H}[\mathbf{u}_i]\mathbf{u}_i^{\mathrm{T}} e_i^3, \quad i \geq 0. \tag{3}$$

Now if

$$\mathrm{H}[\mathbf{u}_i] = \mathrm{diag}\left\{\frac{1}{|\mathbf{u}_{i_1}|}, \frac{1}{|\mathbf{u}_{i_2}|}, \ldots, \frac{1}{|\mathbf{u}_{i_M}|}\right\}, \tag{4}$$

then the update form in (3) reduces to

$$\begin{aligned}
\mathbf{w}_i &= \mathbf{w}_{i-1} + \mu \, \mathrm{diag}\left\{\frac{1}{|\mathbf{u}_{i_1}|}, \frac{1}{|\mathbf{u}_{i_2}|}, \ldots, \frac{1}{|\mathbf{u}_{i_M}|}\right\}\mathbf{u}_i^{\mathrm{T}} e_i^3 \\
&= \mathbf{w}_{i-1} + \mu \, \mathrm{sign}\,[\mathbf{u}_i]^{\mathrm{T}} e_i^3, \quad i \geq 0,
\end{aligned} \tag{5}$$

where M is the filter length. The SRLMF algorithm update recursion in (5) can be regarded as a special case of the general update form in (3) for some matrix data nonlinearity that is implicitly defined by the following relation:

$$\mathrm{sign}\,[\mathbf{u}_i]^{\mathrm{T}} = \mathrm{H}[\mathbf{u}_i]\mathbf{u}_i^{\mathrm{T}}. \tag{6}$$

3. Mean-Square Analysis of the SRLMF Algorithm

We wil assume that the data $\{d_i, \mathbf{u}_i\}$ satisfy the following conditions of the stationary data model [18, 20–24].

(A.1) There exists an optimal weight vector $\mathbf{w}^o$ such that $d_i = \mathbf{u}_i \mathbf{w}^o + v_i$.

(A.2) The noise sequence v_i is independent and identically distributed (i.i.d.) with variance $\sigma_v^2 = \mathrm{E}[|v_i|^2]$ and is independent of $\mathbf{u}_j$ for all i, j.

(A.3) The initial condition $\mathbf{w}_{-1}$ is independent of the zero mean random variables $\{d_i, \mathbf{u}_i, v_i\}$.

(A.4) The regressor covariance matrix is $\mathbf{R} = \mathrm{E}[\mathbf{u}_i^* \mathbf{u}_i] > \mathbf{0}$.

For any adaptive filter of the form in (1), and for any data $\{d_i, \mathbf{u}_i\}$, assuming filter operation in steady-state, the following variance relation holds [18]:

$$\mu \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 \mathrm{g}^2[e_i]\right] = 2\mathrm{E}[e_{a_i}\mathrm{g}[e_i]], \quad \text{as } i \longrightarrow \infty, \tag{7}$$

where

$$\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2\right] = \mathrm{E}\left[\mathbf{u}_i \mathrm{H}[\mathbf{u}_i]\mathbf{u}_i^{\mathrm{T}}\right], \tag{8}$$

$$e_i = e_{a_i} + v_i, \tag{9}$$

and $e_{a_i} = \mathbf{u}_i(\mathbf{w}^o - \mathbf{w}_{i-1})$ is the a priori estimation error. Then $\mathrm{g}[e_i]$ becomes

$$\mathrm{g}[e_i] = e_i^3 = (e_{a_i} + v_i)\left[e_{a_i}^2 + v_i^2 + 2e_{a_i}v_i\right]. \tag{10}$$

By using the fact that e_{a_i} and v_i are independent, we reach at the following expression for the term $\mathrm{E}[e_{a_i}\mathrm{g}[e_i]]$:

$$\mathrm{E}[e_{a_i}\mathrm{g}[e_i]] = 3\sigma_v^2 \mathrm{E}\left[e_{a_i}^2\right] + \mathrm{E}\left[e_{a_i}^4\right]. \tag{11}$$

Ignoring third and higher-order terms of e_{a_i}, then (11) becomes

$$\mathrm{E}[e_{a_i}\mathrm{g}[e_i]] \approx 3\sigma_v^2 \mathrm{E}\left[e_{a_i}^2\right]. \tag{12}$$

To evaluate the term $\mathrm{E}[\|\mathbf{u}_i\|_{\mathrm{H}}^2 \mathrm{g}^2[e_i]]$, we start by noting that

$$\mathrm{g}^2[e_i] = e_{a_i}^6 + 6e_{a_i}^5 v_i + 6e_{a_i} v_i^5 + 15e_{a_i}^4 v_i^2 + 15e_{a_i}^2 v_i^4 \\ + 20e_{a_i}^3 v_i^3 + v_i^6. \tag{13}$$

If we multiply $\mathrm{g}^2[e_i]$ by $\|\mathbf{u}_i\|_{\mathrm{H}}^2$ from the left, use the fact that v_i is independent of both $\mathbf{u}_i$ and e_{a_i}, and again ignoring third and higher-order terms of e_{a_i}, we obtain

$$\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 \mathrm{g}^2[e_i]\right] \approx 6\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i} v_i^5\right] + 15\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}^2 v_i^4\right] \\ + \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 v_i^6\right] \\ \approx 6\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}\right]\mathrm{E}\left[v_i^5\right] + 15\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}^2\right] \\ \times \mathrm{E}\left[v_i^4\right] + \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2\right]\mathrm{E}\left[v_i^6\right] \\ \approx 6\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}\right]\mathrm{E}\left[v_i^5\right] + 15\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}^2\right]\xi_v^4 \\ + \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2\right]\xi_v^6, \tag{14}$$

where $\xi_v^4 = \mathrm{E}[|v_i|^4]$, $\xi_v^6 = \mathrm{E}[|v_i|^6]$ denote the forth and sixth-order moments of v_i, respectively.

From Price's theorem [25] we have

$$\mathrm{E}[x\,\mathrm{sign}(y)] = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma_y}\mathrm{E}[xy], \tag{15}$$

then

$$\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2\right] = \mathrm{E}\left[\mathbf{u}_i\,\mathrm{sign}\,[\mathbf{u}_i]^{\mathrm{T}}\right] = \sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{Tr}(\mathbf{R}). \tag{16}$$

Substituting (16) into (14) we get

$$\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 \mathrm{g}^2[e_i]\right] \approx 6\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}\right]\mathrm{E}\left[v_i^5\right] + 15\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}^2\right]\xi_v^4 \\ + \sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{Tr}(\mathbf{R})\xi_v^6. \tag{17}$$

Substituting (12) and (17) into (7) we get

$$6\sigma_v^2\mathrm{E}\left[e_{a_i}^2\right] = \mu\xi_v^6\sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{Tr}(\mathbf{R}) + 15\mu\xi_v^4\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}^2\right] \\ + 6\mu\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}\right]\mathrm{E}\left[v_i^5\right]. \tag{18}$$

In order to simplify (18) and arrive at an expression for the steady-state EMSE $\zeta = \mathrm{E}[e_{a_i}^2]$, we consider two cases.

(1) Sufficiently Small Step-Sizes. Small step-sizes lead to small values of $\mathrm{E}[e_{a_i}^2]$ and e_{a_i} in steady-state. Therefore, for smaller values of μ, the last two terms in (18) can be ignored, the steady-state EMSE is given by

$$\zeta = \frac{\mu\xi_v^6}{6\sigma_v^2}\sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{Tr}(\mathbf{R}). \tag{19}$$

(2) Separation Principle. For larger values of μ, we resort to the separation assumption, namely, that at steady-state, $\|\mathbf{u}_i\|_{\mathrm{H}}^2$ is independent of e_{a_i}. In this case, the last term in (18) will be zero since e_{a_i} is zero mean, the steady-state EMSE can be shown to be

$$\zeta = \frac{\mu\xi_v^6\sqrt{2/\pi\sigma_u^2}\,\mathrm{Tr}(\mathbf{R})}{\left(6\sigma_v^2 - 15\mu\xi_v^4\sqrt{2/\pi\sigma_u^2}\,\mathrm{Tr}(\mathbf{R})\right)}. \tag{20}$$

4. Convergence Analysis of the SRLMF Algorithm

Convergence analysis of the SRLMF algorithm is much more complicated than that of the LMS algorithm due to existence of the higher order estimation error signal in the coefficient update recursion. We thus make the following assumptions along with (A.2) to make the analysis mathematically more tractable [19–24, 26]:

(A.5) d_i and $\mathbf{u}_i$ are zero-mean, wide-sense stationary, and jointly Gaussian random variables.

(A.6) The input pair $\{d_i, \mathbf{u}_i\}$ is independent of $\{d_j, \mathbf{u}_j\}$ for all i, j.

Subtracting both sides of (5) from $\mathbf{w}^o$ we get

$$\widetilde{\mathbf{w}}_i = \widetilde{\mathbf{w}}_{i-1} + \mu\,\mathrm{sign}\,[\mathbf{u}_i]^{\mathrm{T}}e_i^3, \tag{21}$$

where $\widetilde{\mathbf{w}}_i = \mathbf{w}^o - \mathbf{w}_i$. Taking expectations of both sides of (21) we obtain

$$\mathrm{E}[\widetilde{\mathbf{w}}_i] = \mathrm{E}[\widetilde{\mathbf{w}}_{i-1}] + \mu\mathrm{E}\left[\mathrm{sign}\,[\mathbf{u}_i]^{\mathrm{T}}e_i^3\right]. \tag{22}$$

Using Price's theorem [25], we can conclude that

$$\mathrm{E}\left[\mathrm{sign}\,[\mathbf{u}_i]^{\mathrm{T}}e_i^3\right] = \sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}}e_i^3\right]. \tag{23}$$

Substituting (23) into (22) we get

$$\mathrm{E}[\widetilde{\mathbf{w}}_i] = \mathrm{E}[\widetilde{\mathbf{w}}_{i-1}] + \mu\sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}}e_i^3\right]. \tag{24}$$

The expectation $\mathrm{E}[\mathbf{u}_i^{\mathrm{T}}e_i^3]$ can be simplified using the fact that for zero-mean and jointly Gaussian random variables x_1 and x_2,

$$\mathrm{E}[x_1 x_2^3] = 3\mathrm{E}[x_1 x_2]\mathrm{E}[x_2^2]. \tag{25}$$

Thus, using (25) in conjunction with (A.5), it follows that

$$
\begin{aligned}
\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}} e_i^3\right] &= \mathrm{E}\left[\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}} e_i^3 \mid \widetilde{\mathbf{w}}_{i-1}\right]\right] \\
&= 3\mathrm{E}[\mathrm{E}[e_i^2 \mid \widetilde{\mathbf{w}}_{i-1}]]\mathrm{E}\left[\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}} e_i \mid \widetilde{\mathbf{w}}_{i-1}\right]\right] \\
&= 3\mathrm{E}\left[\sigma_{e|\widetilde{\mathbf{w}}_{i-1}}^2\right]\mathrm{E}\left[\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}} e_i \mid \widetilde{\mathbf{w}}_{i-1}\right]\right],
\end{aligned}
\tag{26}
$$

where

$$
\begin{aligned}
\mathrm{E}\left[\sigma_{e|\widetilde{\mathbf{w}}_{i-1}}^2\right] &= \sigma_e^2 - \mathrm{var}\{\mathrm{E}[e_i \mid \widetilde{\mathbf{w}}_{i-1}]\} \\
&= \sigma_e^2,
\end{aligned}
\tag{27}
$$

and from (9)

$$
\begin{aligned}
\mathrm{E}\left[\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}} e_i \mid \widetilde{\mathbf{w}}_{i-1}\right]\right] &= \mathrm{E}\left[\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}}(\nu_i + \mathbf{u}_i\widetilde{\mathbf{w}}_{i-1}) \mid \widetilde{\mathbf{w}}_{i-1}\right]\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}}\mathbf{u}_i\widetilde{\mathbf{w}}_{i-1} \mid \widetilde{\mathbf{w}}_{i-1}\right]\right] \\
&= \mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}}\mathbf{u}_i\widetilde{\mathbf{w}}_{i-1}\right] \\
&= \mathbf{R}\mathrm{E}[\widetilde{\mathbf{w}}_{i-1}].
\end{aligned}
\tag{28}
$$

Substituting (27) and (28) in (26) yields

$$
\mathrm{E}\left[\mathbf{u}_i^{\mathrm{T}} e_i^3\right] = 3\sigma_e^2\mathbf{R}\mathrm{E}[\widetilde{\mathbf{w}}_{i-1}].
\tag{29}
$$

Substituting (29) into (24) we get

$$
\begin{aligned}
\mathrm{E}[\widetilde{\mathbf{w}}_i] &= \mathrm{E}[\widetilde{\mathbf{w}}_{i-1}] + 3\mu\sqrt{\frac{2}{\pi\sigma_u^2}}\sigma_e^2\mathbf{R}\mathrm{E}[\widetilde{\mathbf{w}}_{i-1}] \\
&= \left[\mathbf{I} + 3\mu\sqrt{\frac{2}{\pi\sigma_u^2}}\sigma_e^2\mathbf{R}\right]\mathrm{E}[\widetilde{\mathbf{w}}_{i-1}].
\end{aligned}
\tag{30}
$$

Ultimately, it is easy to show that the mean behavior of the weight vector, that is $\mathrm{E}[\mathbf{w}_i]$, converges to the optimal weight vector $\mathbf{w}^o$ if μ is bounded by:

$$
0 < \mu < \frac{\sqrt{2\pi\sigma_u^2}}{3\lambda_{\max}\sigma_e^2},
\tag{31}
$$

where $\lambda_{\max}$ represents the maximum eigenvalue of the regressor covariance matrix $\mathbf{R}$. Notice, that there exists the time-varying function σ_e^2 and the regressor variance σ_u^2 in the upper bound for μ. Since σ_e^2 is usually large at the beginning of adaptation processes, we can see that the convergence of the SRLMF algorithm strongly depends on the choice of initial conditions.

5. Tracking Analysis of the SRLMF Algorithm

Here, we assume that the data $\{d_i, \mathbf{u}_i\}$ satisfy the following assumptions of the nonstationary data model [18].

(A.7) There exists a vector $\mathbf{w}_i^o$ such that $d_i = \mathbf{u}_i\mathbf{w}_i^o + \nu_i$.

(A.8) The weight vector varies according to the random-walk model $\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i$, and the sequence $\mathbf{q}_i$ is i.i.d. with covariance matrix $\mathbf{Q}$. Moreover, $\mathbf{q}_i$ is independent of $\{\nu_j, \mathbf{u}_j\}$ for all i, j.

(A.9) The initial conditions $\{\mathbf{w}_{-1}, \mathbf{w}_{-1}^o\}$ are independent of the zero mean random variables $\{d_i, \mathbf{u}_i, \nu_i, \mathbf{q}_i\}$.

In this case, the following variance relation holds [18]:

$$
\mu\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2\mathrm{g}^2[e_i]\right] + \mu^{-1}\mathrm{Tr}(\mathbf{Q}) = 2\mathrm{E}[e_{a_i}\mathrm{g}[e_i]], \quad \text{as } i \longrightarrow \infty.
\tag{32}
$$

Tracking results can be obtained by inspection from the mean-square results as there are only minor differences. Therefore, by substituting (12) and (17) into (32), we get

$$
\begin{aligned}
6\sigma_\nu^2\mathrm{E}\left[e_{a_i}^2\right] &= \mu^{-1}\mathrm{Tr}(\mathbf{Q}) + \mu\xi_\nu^6\sqrt{\frac{2}{\pi\sigma_u^2}}\mathrm{Tr}(\mathbf{R}) \\
&\quad + 15\mu\xi_\nu^4\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}^2\right] \\
&\quad + 6\mu\mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H}}^2 e_{a_i}\right]\mathrm{E}\left[\nu_i^5\right].
\end{aligned}
\tag{33}
$$

We again consider two cases for the evaluation of the tracking EMSE ζ of the SRLMF algorithm.

(1) Sufficiently Small Step-Sizes. Also, here, in this case we get

$$
\zeta = \frac{\mu^{-1}\mathrm{Tr}(\mathbf{Q}) + \mu\xi_\nu^6\sqrt{2/\pi\sigma_u^2}\mathrm{Tr}(\mathbf{R})}{6\sigma_\nu^2}.
\tag{34}
$$

An optimum value of the step-size of the SRLMF algorithm is obtained by minimizing (34) with respect to μ. Setting the derivative of ζ with respect to μ equal to zero gives

$$
\mu_{\mathrm{opt}} = \sqrt{\frac{\mathrm{Tr}(\mathbf{Q})}{\sqrt{2/\pi\sigma_u^2}\mathrm{Tr}(\mathbf{R})\xi_\nu^6}}.
\tag{35}
$$

(2) Separation Principle. Similarly here as it was done for the derivation of (20), we obtain the following:

$$
\zeta = \frac{\mu^{-1}\mathrm{Tr}(\mathbf{Q}) + \mu\xi_\nu^6\sqrt{2/\pi\sigma_u^2}\mathrm{Tr}(\mathbf{R})}{\left(6\sigma_\nu^2 - 15\mu\xi_\nu^4\sqrt{2/\pi\sigma_u^2}\mathrm{Tr}(\mathbf{R})\right)},
\tag{36}
$$

and eventually the optimum step-size of the SRLMF algorithm is given by

$$
\begin{aligned}
\mu_{\mathrm{opt}} &= \sqrt{\mathrm{Tr}(\mathbf{Q})\left[\frac{225\left(\xi_\nu^4\right)^2\mathrm{Tr}(\mathbf{Q})}{36\left(\sigma_\nu^2\right)^2\left(\xi_\nu^6\right)^2} + \frac{1}{\sqrt{2/\pi\sigma_u^2}\mathrm{Tr}(\mathbf{R})\xi_\nu^6}\right]} \\
&\quad - \frac{15\xi_\nu^4}{6\sigma_\nu^2\xi_\nu^6}\mathrm{Tr}(\mathbf{Q}).
\end{aligned}
\tag{37}
$$

6. Transient Analysis of the SRLMF Algorithm

Here, we will assume that the data $\{d_i, \mathbf{u}_i\}$ satisfy the conditions of the stationary data model described in Section 3.

6.1. Weighted Energy-Conservation Relation

Theorem 1. *For any adaptive filter of the form* (1), *any positive-definite Hermitian matrix* Σ, *and for any data* $\{d_i, \mathbf{u}_i\}$, *it holds that [18]:*

$$\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 + \left| e_{a_i}^{\mathrm{H\Sigma}} \right|^2 = \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \left| e_{p_i}^{\mathrm{H\Sigma}} \right|^2,$$
(38)

where $e_{a_i}^{\mathrm{H\Sigma}} \mathbf{u}_i H[\mathbf{u}_i] \Sigma \widetilde{\mathbf{w}}_{i-1}$, $e_{p_i}^{\mathrm{H\Sigma}} \mathbf{u}_i H[\mathbf{u}_i] \Sigma \widetilde{\mathbf{w}}_i$, *and* $\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 = \mathbf{u}_i (H[\mathbf{u}_i] \Sigma H[\mathbf{u}_i]) \mathbf{u}_i^*$.

Proof. Let us consider the adaptive filter updates of the generic form given in (1). Subtracting both sides of (1) from $\mathbf{w}^o$, we get

$$\widetilde{\mathbf{w}}_i = \widetilde{\mathbf{w}}_{i-1} - \mu H[\mathbf{u}_i] \mathbf{u}_i^* g[e_i].$$
(39)

If we multiply both sides of (39) by $\mathbf{u}_i H[\mathbf{u}_i] \Sigma$ from the left, we get

$$e_{p_i}^{\mathrm{H\Sigma}} = e_{a_i}^{\mathrm{H\Sigma}} - \mu \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g[e_i].$$
(40)

Two cases can be considered here.

Case 1 ($\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 = 0$). In this case, $\widetilde{\mathbf{w}}_i = \widetilde{\mathbf{w}}_{i-1}$ and $e_{a_i}^{\mathrm{H\Sigma}} = e_{p_i}^{\mathrm{H\Sigma}}$ so that $\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 = \|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2$ and $|e_{a_i}^{\mathrm{H\Sigma}}|^2 = |e_{p_i}^{\mathrm{H\Sigma}}|^2$.

Case 2 ($\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \neq 0$). In this case, we use (40) to solve for $g[e_i]$,

$$g[e_i] = \frac{1}{\mu \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} \left(e_{a_i}^{\mathrm{H\Sigma}} - e_{p_i}^{\mathrm{H\Sigma}} \right).$$
(41)

Substituting (41) into (39), we get

$$\widetilde{\mathbf{w}}_i = \widetilde{\mathbf{w}}_{i-1} - \frac{H[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} \left(e_{a_i}^{\mathrm{H\Sigma}} - e_{p_i}^{\mathrm{H\Sigma}} \right).$$
(42)

Expression (42) can be rearranged as

$$\widetilde{\mathbf{w}}_i + \frac{H[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} e_{a_i}^{\mathrm{H\Sigma}} = \widetilde{\mathbf{w}}_{i-1} + \frac{H[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} e_{p_i}^{\mathrm{H\Sigma}}.$$
(43)

Evaluating the energies of both sides of (43) results in

$$\left\| \widetilde{\mathbf{w}}_i + \frac{H[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} e_{a_i}^{\mathrm{H\Sigma}} \right\|_{\Sigma}^2 = \left\| \widetilde{\mathbf{w}}_{i-1} + \frac{H[\mathbf{u}_i] \mathbf{u}_i^*}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} e_{p_i}^{\mathrm{H\Sigma}} \right\|_{\Sigma}^2.$$
(44)

After a straightforward calculation, the following weighted energy-conservation results:

$$\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 + \frac{1}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} \left| e_{a_i}^{\mathrm{H\Sigma}} \right|^2 = \|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \frac{1}{\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2} \left| e_{p_i}^{\mathrm{H\Sigma}} \right|^2.$$
(45)

The weighted energy-conservation relation in (45) can also be written as

$$\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 + \left| e_{a_i}^{\mathrm{H\Sigma}} \right|^2 = \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \left| e_{p_i}^{\mathrm{H\Sigma}} \right|^2.$$
(46)

$\square$

6.2. Weighted Variance Relation.

Here, the weighted variance relation presented in [18] has been extended in order to derive expressions for the MSE and the MSD of the SRLMF algorithm during the transient phase.

Theorem 2. *For any adaptive filter of the form* (1), *any positive-definite Hermitian matrix* Σ, *and for any data* $\{d_i, \mathbf{u}_i\}$, *it holds that*

$$\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 \right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 \right] + \mu^2 \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \left| g[e_i] \right|^2 \right] - 2\mu \operatorname{Re}\left(\mathrm{E}\left[e_{a_i}^{\mathrm{H\Sigma}*} g[e_i] \right] \right), \quad as\ i \longrightarrow \infty.$$
(47)

Similarly, for real-valued data, the above weighted variance relation becomes

$$\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 \right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 \right] + \mu^2 \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g^2[e_i] \right] - 2\mu \mathrm{E}\left[e_{a_i}^{\mathrm{H\Sigma}} g[e_i] \right], \quad as\ i \longrightarrow \infty.$$
(48)

Proof. Squaring both sides of (40), we get

$$\left| e_{p_i}^{\mathrm{H\Sigma}} \right|^2 = \left| e_{a_i}^{\mathrm{H\Sigma}} - \mu \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g[e_i] \right|^2.$$
(49)

For compactness of notation let us omit the argument of g so that (49) looks like

$$\left| e_{p_i}^{\mathrm{H\Sigma}} \right|^2 = \left| e_{a_i}^{\mathrm{H\Sigma}} \right|^2 + \mu^2 \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^4 |g|^2 - \mu e_{a_i}^{\mathrm{H\Sigma}} \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g^* - \mu e_{a_i}^{\mathrm{H\Sigma}*} \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g.$$
(50)

Substituting (50) into (46), we get

$$\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 = \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \mu^2 \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^4 |g|^2 - \mu e_{a_i}^{\mathrm{H\Sigma}} \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g^* - \mu e_{a_i}^{\mathrm{H\Sigma}*} \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g.$$
(51)

Dividing both sides of (51) by $\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2$ (of course here $\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \neq 0$) we get

$$\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 = \|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 + \mu^2 \|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 |g|^2 - \mu e_{a_i}^{\mathrm{H\Sigma}} g^* - \mu e_{a_i}^{\mathrm{H\Sigma}*} g.$$
(52)

Taking expectations of both sides of (52), we obtain

$$\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 \right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 \right] + \mu^2 \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \left| g[e_i] \right|^2 \right] - \mu \mathrm{E}\left[e_{a_i}^{\mathrm{H\Sigma}} g[e_i]^* + e_{a_i}^{\mathrm{H\Sigma}*} g[e_i] \right],$$
(53)

or in the following format:

$$\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 \right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 \right] + \mu^2 \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 \left| g[e_i] \right|^2 \right] - 2\mu \operatorname{Re}\left(\mathrm{E}\left[e_{a_i}^{\mathrm{H\Sigma}*} g[e_i] \right] \right), \quad as\ i \longrightarrow \infty.$$
(54)

For real-valued data, the weighted variance relation in (54) becomes

$$\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\Sigma}^2 \right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\Sigma}^2 \right] + \mu^2 \mathrm{E}\left[\|\mathbf{u}_i\|_{\mathrm{H\Sigma H}}^2 g^2[e_i] \right] - 2\mu \mathrm{E}\left[e_{a_i}^{\mathrm{H\Sigma}} g[e_i] \right], \quad as\ i \longrightarrow \infty.$$
(55)

$\square$

The transient analysis of the class of filters in (1) is more challenging due to the presence of the error nonlinearity. Nevertheless, by using some approximations, the analysis can be carried out to provide some useful insights about the performance of the SRLMF algorithm.

To start, the expectations $\mathrm{E}[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}g^2[e_i]]$ and $\mathrm{E}[e^{\mathrm{H}\Sigma}_{a_i}g[e_i]]$ are evaluated in the ensuing analysis in terms of the weighted norm of $\widetilde{\mathbf{w}}_{i-1}$. Since these expectations are involved mathematically we will rely on the following assumption in order to facilitate their evaluation [18].

(A.10) The a priori estimation errors $\{e_{a_i}, e^{\mathrm{H}\Sigma}_{a_i}\}$ are jointly circular Gaussian.

Evaluation of $E[e^{H\Sigma}_{a_i}g[e_i]]$. From Price's theorem, if x and y are jointly Gaussian random variables that are independent from a third random variable z, then it holds that [25]:

$$\mathrm{E}[xg(y+z)] = \frac{\mathrm{E}[xy]}{\mathrm{E}[y^2]}\mathrm{E}[yg(y+z)]. \qquad (56)$$

Applying this result to the term $\mathrm{E}[e^{\mathrm{H}\Sigma}_{a_i}g[e_i]]$, and using (9), we get

$$\mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}g[e_i]\right] = \mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}g[e_{a_i}+v_i]\right]$$
$$= \mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}e_{a_i}\right]\left[\frac{\mathrm{E}[e_{a_i}g[e_i]]}{\mathrm{E}\left[e^2_{a_i}\right]}\right]. \qquad (57)$$

In view of the assumption (A.10), the expectation $\mathrm{E}[e_{a_i}g[e_i]]$ depends on e_{a_i} only through its second moment, $\mathrm{E}[e^2_{a_i}]$. Therefore, we can define the following function of $\mathrm{E}[e^2_{a_i}]$:

$$\mathcal{Z}_1 = \frac{\mathrm{E}[e_{a_i}g[e_i]]}{\mathrm{E}\left[e^2_{a_i}\right]}. \qquad (58)$$

For the SRLMF algorithm, $g[e_i] = e^3_i$, therefore

$$\mathrm{E}[e_{a_i}g[e_i]] = \mathrm{E}\left[e_{a_i}(e_{a_i}+v_i)^3\right]$$
$$= \mathrm{E}\left[e^4_{a_i}+3e^3_{a_i}v_i+3e^2_{a_i}v^2_i+v^3_ie_{a_i}\right]. \qquad (59)$$

Now since e_{a_i} and v_i are zero mean Gaussian and independent random variables with variances $\mathrm{E}[e^2_{a_i}]$ and σ^2_v, respectively, we obtain

$$\mathrm{E}[e_{a_i}g[e_i]] = \mathrm{E}\left[e^4_{a_i}\right]+3\sigma^2_v\mathrm{E}\left[e^2_{a_i}\right]. \qquad (60)$$

By using the fact that for circular Gaussian e_{a_i} it holds that $\mathrm{E}[e^4_{a_i}] = 3\mathrm{E}[e^2_{a_i}]^2$, we get

$$\mathrm{E}[e_{a_i}g[e_i]] = 3\mathrm{E}\left[e^2_{a_i}\right]^2+3\sigma^2_v\mathrm{E}\left[e^2_{a_i}\right]$$
$$= 3\mathrm{E}\left[e^2_{a_i}\right]\left[\mathrm{E}\left[e^2_{a_i}\right]+\sigma^2_v\right]. \qquad (61)$$

Substituting (61) into (58), we get

$$\mathcal{Z}_1 = 3\left[\mathrm{E}\left[e^2_{a_i}\right]+\sigma^2_v\right]. \qquad (62)$$

The expression for $\mathcal{Z}_1$ is related to the desired term $\mathrm{E}[e^{\mathrm{H}\Sigma}_{a_i}g[e_i]]$ through the equality

$$\mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}g[e_i]\right] = \mathcal{Z}_1\mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}e_{a_i}\right]. \qquad (63)$$

Evaluation of $E[\|\mathbf{u}_i\|^2_{H\Sigma H}g^2[e_i]]$. In order to facilitate the evaluation of the term $\mathrm{E}[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}g^2[e_i]]$ we use the separation principle, namely, we assume that the filter is long enough so that the following assumption holds [18].

(A.11) $\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}$ is independent of e_i.

Therefore,

$$\mathrm{E}\left[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}g^2[e_i]\right] = \left(\mathrm{E}\left[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}\right]\right)\left(\mathrm{E}[g^2[e_i]]\right). \qquad (64)$$

Since e_{a_i} is Gaussian and independent of the noise, the expectation $\mathrm{E}[g^2[e_i]]$ depends on e_{a_i} through its second moment only. Therefore, we can define the following function of $\mathrm{E}[e^2_{a_i}]$:

$$\mathcal{Z}_2 = \mathrm{E}[g^2[e_i]]. \qquad (65)$$

For the SRLMF algorithm, $g[e_i] = e^3_i$. Since e_{a_i} and v_i are zero mean Gaussian and independent random variables with variances $\mathrm{E}[e^2_{a_i}]$ and σ^2_v, we have $\sigma^2_e = \mathrm{E}[e^2_i] = \mathrm{E}[e^2_{a_i}]+\sigma^2_v$. Moreover from [18], $\mathrm{E}[e^6_i] = 15\sigma^6_e$. Thus

$$\mathcal{Z}_2 = \mathrm{E}\left[e^6_i\right]$$
$$= 15\sigma^6_e$$
$$= 15(\sigma^2_e)^3$$
$$= 15\left(\mathrm{E}\left[e^2_{a_i}\right]+\sigma^2_v\right)^3$$
$$= 15\left(\mathrm{E}\left[e^2_{a_i}\right]\right)^3+45\sigma^2_v\left(\mathrm{E}\left[e^2_{a_i}\right]\right)^2+45\xi^4_v\mathrm{E}\left[e^2_{a_i}\right]+15\xi^6_v. \qquad (66)$$

The expression for $\mathcal{Z}_2$ is related to the desired term $\mathrm{E}[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}g^2[e_i]]$ through the equality

$$\mathrm{E}\left[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}g^2[e_i]\right] = \mathcal{Z}_2\mathrm{E}\left[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}\right]$$
$$= \mathcal{Z}_2\mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|^2_\Sigma\right]. \qquad (67)$$

Since

$$\mathrm{E}\left[\|\mathbf{u}_i\|^2_{\mathrm{H}\Sigma\mathrm{H}}\right] = \mathrm{E}\left[\mathbf{u}_i\mathrm{H}[\mathbf{u}_i]\Sigma\mathrm{H}[\mathbf{u}_i]\mathbf{u}^\mathrm{T}_i\right]$$
$$= \mathrm{E}\left[\mathrm{sign}[\mathbf{u}_i]\Sigma\,\mathrm{sign}[\mathbf{u}_i]^\mathrm{T}\right] \qquad (68)$$
$$= \mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|^2_\Sigma\right].$$

Substituting (63) and (67) into (55), we get

$$\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|^2_\Sigma\right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|^2_\Sigma\right]+\mu^2\mathcal{Z}_2\mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|^2_\Sigma\right]$$
$$-2\mu\mathcal{Z}_1\mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}e_{a_i}\right]. \qquad (69)$$

Independence Assumption. If we assume that the regressor sequence $\{\mathbf{u}_i\}$ is i.i.d. then

$$\mathrm{E}\left[e^{\mathrm{H}\Sigma}_{a_i}e_{a_i}\right] = \mathrm{E}\left[\widetilde{\mathbf{w}}^\mathrm{T}_{i-1}\Sigma\mathrm{H}[\mathbf{u}_i]\mathbf{u}^\mathrm{T}_i\mathbf{u}_i\widetilde{\mathbf{w}}_{i-1}\right]$$
$$= \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|^2_{\Sigma\mathrm{H}\mathbf{u}^\mathrm{T}_i\mathbf{u}_i}\right]. \qquad (70)$$

In this way, the terms $\{E[e_{a_i}^{H\Sigma}e_{a_i}], Z_1, Z_2\}$ become all functions of $\widetilde{\mathbf{w}}_{i-1}$. Therefore, (69) becomes

$$
\begin{aligned}
E\left[||\widetilde{\mathbf{w}}_i||_\Sigma^2\right] &= E\left[||\widetilde{\mathbf{w}}_{i-1}||_\Sigma^2\right] + \mu^2 Z_2 E\left[||\text{sign}[\mathbf{u}_i]||_\Sigma^2\right] \\
&\quad - 2\mu Z_1 E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\Sigma H \mathbf{u}_i^T \mathbf{u}_i}^2\right] \\
&= E\left[||\widetilde{\mathbf{w}}_{i-1}||_\Sigma^2\right] + \mu^2 Z_2 E\left[||\text{sign}[\mathbf{u}_i]||_\Sigma^2\right] \\
&\quad - 2\mu Z_1 E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\Sigma \,\text{sign}\,[\mathbf{u}_i]^T \mathbf{u}_i}^2\right] \\
&= E\left[||\widetilde{\mathbf{w}}_{i-1}||_\Sigma^2\right] + \mu^2 Z_2 E\left[||\text{sign}[\mathbf{u}_i]||_\Sigma^2\right] \\
&\quad - \sqrt{\frac{8}{\pi\sigma_u^2}} \mu Z_1 E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\Sigma R}^2\right].
\end{aligned}
\tag{71}
$$

We thus find that studying the transient behavior of the SRLMF algorithm in effect has reduced to evaluating the functions Z_1 and Z_2 and studying the resulting variance relation (71). Let us now illustrate the application of the above results for white and correlated input data.

White Input Data. For white input data $\mathbf{R}$ is diagonal, say $\mathbf{R} = \sigma_u^2 \mathbf{I}$. Therefore, if we select $\Sigma = \mathbf{I}$, the variance relation (71) becomes

$$
\begin{aligned}
E\left[||\widetilde{\mathbf{w}}_i||^2\right] &= E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + \mu^2 Z_2 E\left[||\text{sign}[\mathbf{u}_i]||^2\right] \\
&\quad - \sqrt{\frac{8\sigma_u^2}{\pi}} \mu Z_1 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right].
\end{aligned}
\tag{72}
$$

Now since

$$
\begin{aligned}
e_{a_i}^2 &= \widetilde{\mathbf{w}}_{i-1}^T \mathbf{u}_i^T \mathbf{u}_i \widetilde{\mathbf{w}}_{i-1} \\
&= ||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{u}_i^T \mathbf{u}_i}^2.
\end{aligned}
\tag{73}
$$

Substituting (73) into (66), we get

$$
\begin{aligned}
Z_2 &= 15\left(E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{u}_i^T \mathbf{u}_i}^2\right]\right)^3 + 45\sigma_v^2\left(E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{u}_i^T \mathbf{u}_i}^2\right]\right)^2 \\
&\quad + 45\xi_v^4 E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{u}_i^T \mathbf{u}_i}^2\right] + 15\xi_v^6 \\
&= 15\left(E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{R}}^2\right]\right)^3 + 45\sigma_v^2\left(E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{R}}^2\right]\right)^2 \\
&\quad + 45\xi_v^4 E\left[||\widetilde{\mathbf{w}}_{i-1}||_{\mathbf{R}}^2\right] + 15\xi_v^6 \\
&= 15\left(\sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^3 + 45\sigma_v^2\left(\sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^2 \\
&\quad + 45\xi_v^4 \sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + 15\xi_v^6.
\end{aligned}
\tag{74}
$$

Similarly by substituting (73) into (62), we get

$$
Z_1 = 3\left(\sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + \sigma_v^2\right).
\tag{75}
$$

Substituting (74) and (75) into (72), we get

$$
\begin{aligned}
E\left[||\widetilde{\mathbf{w}}_i||^2\right] &= E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + \mu^2\bigg[15\left(\sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^3 \\
&\quad + 45\sigma_v^2\left(\sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^2 \\
&\quad + 45\xi_v^4 \sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + 15\xi_v^6\bigg] \\
&\quad \times E\left[||\text{sign}[\mathbf{u}_i]||^2\right] - 3\sqrt{\frac{8\sigma_u^2}{\pi}}\mu \\
&\quad \times \left(\sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + \sigma_v^2\right) E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right].
\end{aligned}
\tag{76}
$$

Since $E[||\,\text{sign}[\mathbf{u}_i]\,||^2] = M$, the recursion in (76) becomes

$$
\begin{aligned}
E\left[||\widetilde{\mathbf{w}}_i||^2\right] &= E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + 15\mu^2 M \sigma_u^6 \left(E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^3 \\
&\quad + 45\mu^2 M \sigma_v^2 \sigma_u^4 \left(E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^2 \\
&\quad + 45\mu^2 M \xi_v^4 \sigma_u^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + 15\mu^2 M \xi_v^6 \\
&\quad - 6\sqrt{\frac{2\sigma_u^2}{\pi}}\mu\sigma_u^2 \left(E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^2 \\
&\quad - 6\sqrt{\frac{2\sigma_u^2}{\pi}}\mu\sigma_v^2 E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] \\
&= f E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + 15\mu^2 M \xi_v^6,
\end{aligned}
\tag{77}
$$

where

$$
\begin{aligned}
f &= 1 + 3\mu\left(15\mu M \sigma_u^2 \xi_v^4 - 2\sqrt{\frac{2\sigma_u^2}{\pi}}\sigma_v^2\right) \\
&\quad + 3\mu\sigma_u^2\left(15\mu M \sigma_u^2 \sigma_v^2 - 2\sqrt{\frac{2\sigma_u^2}{\pi}}\right) E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] \\
&\quad + 15\mu^2 M \sigma_u^6 \left(E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right]\right)^2.
\end{aligned}
\tag{78}
$$

We see that the transient behavior of the SRLMF algorithm is described by a nonlinear recursion in $E[||\widetilde{\mathbf{w}}_i||^2]$ due to the presence of the factor $E[||\widetilde{\mathbf{w}}_{i-1}||^2]$ inside f.

Correlated Input Data. For uncorrelated data, the variance relation (72) shows that only unweighted norms of $\widetilde{\mathbf{w}}_i$ and $\widetilde{\mathbf{w}}_{i-1}$ appear on both sides of the equation. However, for correlated data, different weighing matrices will appear on both sides of (72).

If $\Sigma = \mathbf{I}$ in (71), we get

$$
\begin{aligned}
E\left[||\widetilde{\mathbf{w}}_i||^2\right] &= E\left[||\widetilde{\mathbf{w}}_{i-1}||^2\right] + \mu^2 Z_2 E\left[||\text{sign}[\mathbf{u}_i]||^2\right] \\
&\quad - \sqrt{\frac{8}{\pi\sigma_u^2}}\mu Z_1 E\left[||\widetilde{\mathbf{w}}_{\mathbf{i}-1}||_{\mathbf{R}}^2\right].
\end{aligned}
\tag{79}
$$

If $\Sigma = \mathbf{R}$ in (71), we get

$$
\begin{aligned}
\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}}^2\right] &= \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}}^2\right] + \mu^2 \mathcal{Z}_2 \mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|_{\mathbf{R}}^2\right] \\
&\quad - \sqrt{\frac{8}{\pi\sigma_u^2}}\mu\mathcal{Z}_1 \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}^2}^2\right].
\end{aligned}
\tag{80}
$$

Similarly if $\Sigma = \mathbf{R}^{M-1}$ in (71), we get

$$
\begin{aligned}
\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^{M-1}}^2\right] &= \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}^{M-1}}^2\right] + \mu^2 \mathcal{Z}_2 \mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|_{\mathbf{R}^{M-1}}^2\right] \\
&\quad - \sqrt{\frac{8}{\pi\sigma_u^2}}\mu\mathcal{Z}_1 \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}^M}^2\right].
\end{aligned}
\tag{81}
$$

The term $\mathrm{E}[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^M}^2]$ can be inferred from the prior weighting factors

$$
\left\{\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|^2\right], \mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}}^2\right], \mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^2}^2\right], \ldots, \mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^{M-1}}^2\right]\right\},
\tag{82}
$$

by expressing $\mathbf{R}^M$ as a linear combination of its lower-order powers using the Cayley-Hamilton theorem. Thus let $p(x) = \det(x\mathrm{I} - \mathbf{R})$ denote the characteristic polynomial of $\mathbf{R}$, say

$$
p(x) = x^M + p_{M-1}x^{M-1} + p_{M-2}x^{M-2} + \cdots + p_1 x + p_0.
\tag{83}
$$

Then we know that [18]:

$$
\mathbf{R}^M = -p_{M-1}\mathbf{R}^{M-1} - p_{M-2}\mathbf{R}^{M-2} - \cdots - p_1\mathbf{R} - p_0\mathrm{I}.
\tag{84}
$$

Using this fact, we have

$$
\begin{aligned}
\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^M}^2\right] &= -p_0\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|^2\right] - p_1\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}}^2\right] \\
&\quad - \cdots \\
&\quad - p_{M-1}\mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^{M-1}}^2\right].
\end{aligned}
\tag{85}
$$

We can collect the above results into a compact vector notation by writing (79)–(81) as

$$
\mathcal{W}_i = \mathcal{F}\,\mathcal{W}_{i-1} + \mu^2\mathcal{Z}_2\mathcal{Y},
\tag{86}
$$

where the $M \times 1$ vectors $\{\mathcal{W}_i, \mathcal{Y}\}$ are given by

$$
\mathcal{W}_i = \begin{bmatrix} \mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|^2\right] \\ \mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}}^2\right] \\ \vdots \\ \mathrm{E}\left[\|\widetilde{\mathbf{w}}_i\|_{\mathbf{R}^{M-1}}^2\right] \end{bmatrix}, \qquad \mathcal{Y} = \begin{bmatrix} \mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|^2\right] \\ \mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|_{\mathbf{R}}^2\right] \\ \vdots \\ \mathrm{E}\left[\|\mathrm{sign}[\mathbf{u}_i]\|_{\mathbf{R}^{M-1}}^2\right] \end{bmatrix},
\tag{87}
$$

and the $M \times M$ coefficient matrix $\mathcal{F}$ is given by

$$
\mathcal{F} = \begin{bmatrix}
1 & -\sqrt{\frac{8}{\pi\sigma_u^2}}\mu\mathcal{Z}_1 & & & \\
0 & 1 & -\sqrt{\frac{8}{\pi\sigma_u^2}}\mu\mathcal{Z}_1 & & \\
0 & 0 & 1 & -\sqrt{\frac{8}{\pi\sigma_u^2}}\mu\mathcal{Z}_1 & \\
\vdots & & & & \\
0 & 0 & & 1 & -\sqrt{\frac{8}{\pi\sigma_u^2}}\mu\mathcal{Z}_1 \\
\sqrt{\frac{8}{\pi\sigma_u^2}}\mu p_0\mathcal{Z}_1 & \sqrt{\frac{8}{\pi\sigma_u^2}}\mu p_1\mathcal{Z}_1 & \cdots & \sqrt{\frac{8}{\pi\sigma_u^2}}\mu p_{M-2}\mathcal{Z}_1 & 1 + \sqrt{\frac{8}{\pi\sigma_u^2}}\mu p_{M-1}\mathcal{Z}_1
\end{bmatrix}.
\tag{88}
$$

As can be seen from (86), the transient behavior of the SRLMF algorithm is described by an M-dimensional state-space recursion as opposed to one-dimensional in the white input case (72).

We know that, the mean-square error is defined as

$$
\mathrm{MSE} \triangleq \lim_{i\to\infty}\mathrm{E}\left[|e_i|^2\right],
\tag{89}
$$

and the excess mean-square error is defined as

$$
\mathrm{EMSE} \triangleq \lim_{i\to\infty}\mathrm{E}\left[|e_{a_i}|^2\right],
\tag{90}
$$

where

$$
\mathrm{E}\left[|e_{a_i}|^2\right] = \mathrm{E}\left[\|\widetilde{\mathbf{w}}_{i-1}\|_{\mathbf{R}}^2\right].
\tag{91}
$$

TABLE 1: Computational load per iteration for LMF and SRLMF algorithms when data is real.

Algorithm	+	×	Sign
LMF	$2M$	$2M + 3$	
SRLMF	$2M$	$2M + 2$	1

TABLE 2: Computational load per iteration for LMF and SRLMF algorithms when data is complex.

Algorithm	+	×	Sign
LMF	$8M + 1$	$8M + 5$	
SRLMF	$6M + 1$	$6M + 3$	2

The evolution of $E[|e_{a_i}|^2]$ is described by the second entry of the state vector $\mathcal{W}_i$ in (86). The resulting learning curve of the filter is $E[|e_i|^2] = \sigma_v^2 + E[|e_{a_i}|^2]$.

We know that the mean-square deviation is defined as

$$\text{MSD} \triangleq \lim_{i \to \infty} E\left[\|\widetilde{\mathbf{w}}_i\|^2\right]. \tag{92}$$

The evolution of $E[\|\widetilde{\mathbf{w}}_i\|^2]$ is described by the first entry of the state vector $\mathcal{W}_i$ in (86).

7. Computational Load

Finally, the computational complexity of the LMF and SRLMF algorithms is discussed in this section. Tables 1 and 2 detail the estimated computational load per iteration for LMF and SRLMF algorithms, respectively, for real- and complex-valued data in terms of the number of real additions (+), real multiplications (×), and comparisons with zero (or sign evaluations). We know that one complex multiplication requires four real multiplications and two real additions, while one complex addition requires two real additions.

As can be seen from these two tables, the computational complexity of the SRLMF algorithm becomes more interesting when the data is complex-valued. The case of fading channels in mobile communications is a good example where this scenario can bring drastic improvement in complexity of the SRLMF algorithm over the LMF algorithm.

8. Simulation Results

First, the performance analysis of the LMF and the SRLMF algorithms is investigated in an unknown system identification setup with $\mathbf{w}^o = [0.227 \quad 0.460 \quad 0.688 \quad 0.460 \quad 0.227]^{\text{T}}$ as far as convergence, steady-state and transient behaviors are concerned. Figure 1 depicts the convergence behavior of the two algorithms for a signal to noise ratio (SNR) of 10 dB in a uniform environment. This figure shows almost identical performance for the two algorithms; no deterioration has occurred to the SRLMF algorithm.

Second, in order to validate the theoretical findings, extensive simulations are carried out for different scenarios. While Figures 2–4 are for the case of the steady-state EMSE of the SRLMF algorithm in a stationary environment, Figure 5 is for the case of the tracking EMSE in a nonstationary
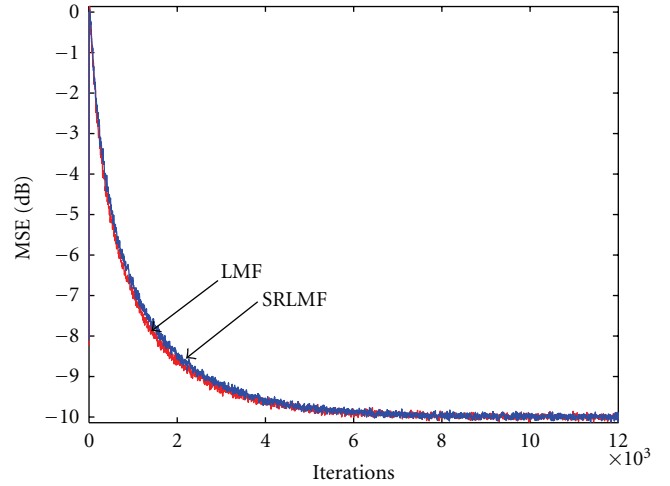

FIGURE 1: Comparison of the MSE learning curves of LMF and SRLMF algorithms in a uniform noise environment with SNR = 10 dB.
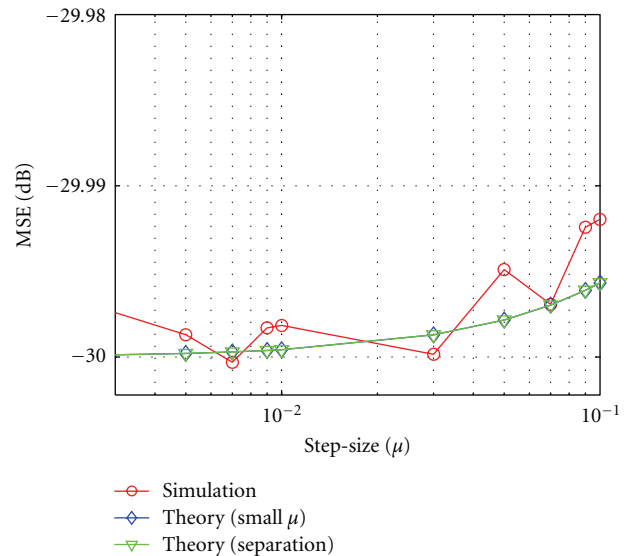

FIGURE 2: Theoretical and simulated MSE learning curves of the SRLMF algorithm using white Gaussian regressors with shift structure with SNR = 30 dB.

environment. In all of these figures the MSE is plotted versus the step-size μ with a SNR = 30 dB.

In the case of Figure 2, the regressors, with shift structure, are generated by feeding a unit-variance white process into a tapped delay line. However, in Figure 3, the regressors, with shift structure, are generated by passing correlated data into a tapped delay line. Here, the correlated data are obtained by passing a unit-variance i.i.d. Gaussian data through a first-order autoregressive model with transfer function $\sqrt{1 - a^2}/(1 - az^{-1})$ and $a = 0.8$. To further test the validity of the results, Gaussian regressors with an eigenvalue spread of five without a shift structure are used, this is depicted in Figure 4. As it can be seen from these figures, the simulation results match very well the theoretical results ((19) and (20)).
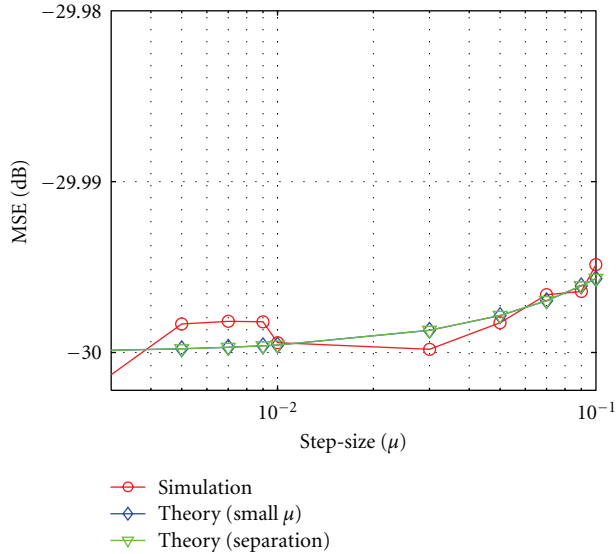
FIGURE 3: Theoretical and simulated MSE learning curves of the SRLMF algorithm using correlated Gaussian regressors with shift structure with SNR = 30 dB.
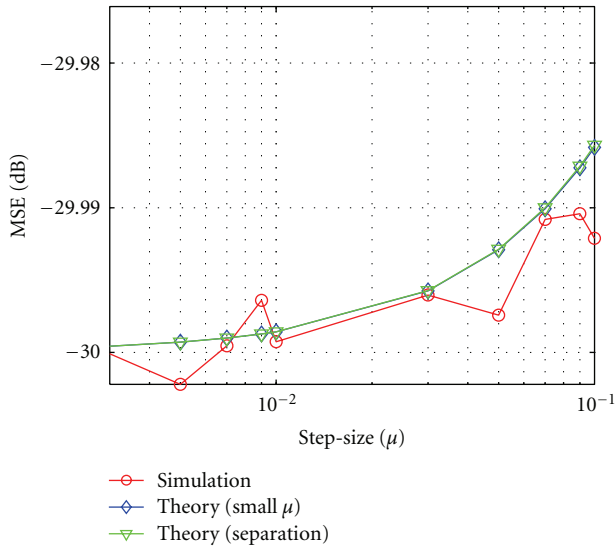

FIGURE 4: Theoretical and simulated MSE learning curves of the SRLMF algorithm using Gaussian regressors with an eigenvalue spread = 5 without shift structure with SNR = 30 dB.
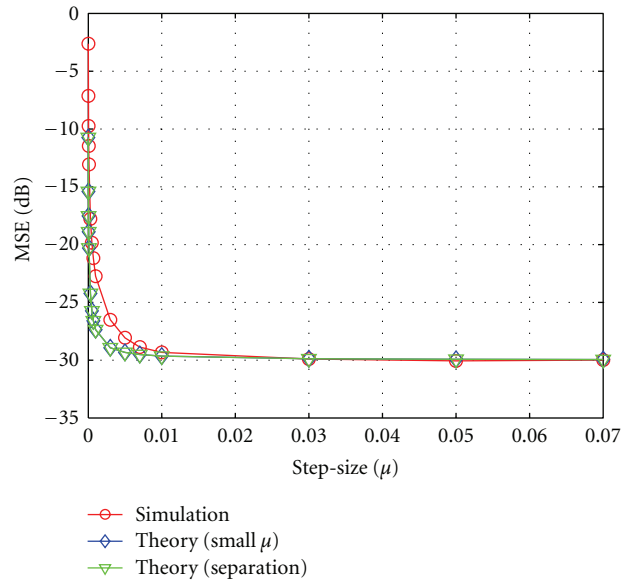

FIGURE 5: Theoretical and simulated MSE learning curves of the SRLMF algorithm for a random-walk channel with SNR = 30 dB.
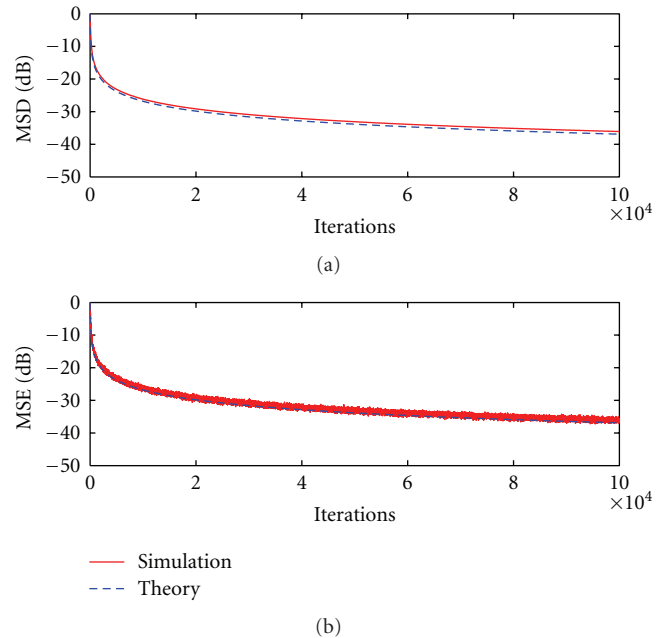

FIGURE 6: Theoretical and simulated MSD (a) and MSE (b) learning curves of the SRLMF algorithm using white Gaussian regressors with SNR = 50 dB.

Third, to further validate the theoretical results in a tracking scenario, the results of Figure 5 depicts this behavior. Here, the random-walk channel behaves according to

$$\mathbf{w}_i^o = \mathbf{w}_{i-1}^o + \mathbf{q}_i, \tag{93}$$

where $\mathbf{q}_i$ is a Gaussian sequence with zero mean and variance $\sigma_q^2 = 10^{-9}$ and $\mathbf{w}_{-1}^o = \mathbf{w}^o$. As observed from Figure 5, the simulation results corroborate closely the theoretical results ((34) and (36)).

Finally, we examine the transient behavior of the SRLMF algorithm for the case of Gaussian data. Let us consider a real-valued regression sequence $\{\mathbf{u}_i\}$ with covariance matrix $\mathbf{R}$ whose eigenvalue spread we set at $\rho = 5$. Let the SNR be 50 dB and the step-size is fixed at $\mu = 0.01$.

The results in Figures 6 and 7 show the theoretical and simulated MSD and MSE learning curves of the SRLMF algorithm using white Gaussian regressors and Gaussian regressors with an eigenvalue spread equal to 5. The theoretical values are obtained by using the expression (86). As can be seen here, There is an excellent match between the theoretical and simulated results.
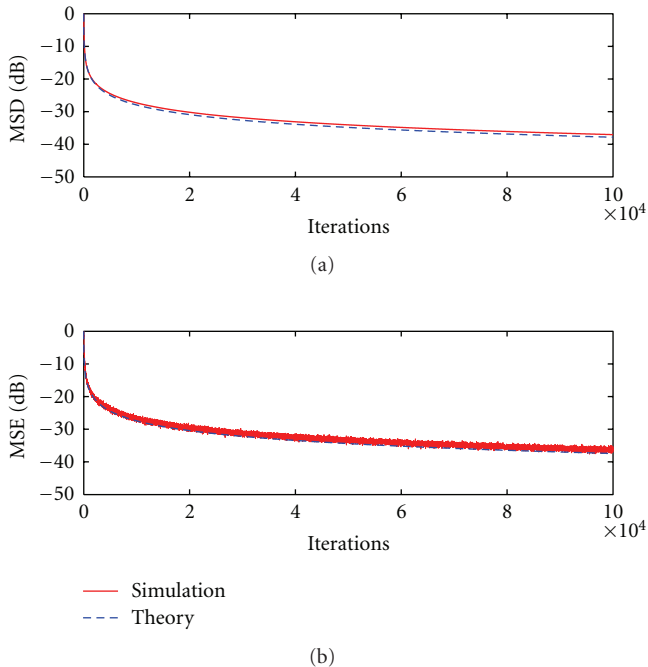
FIGURE 7: Theoretical and simulated MSD (a) and MSE (b) learning curves of the SRLMF algorithm using Gaussian regressors with an eigenvalue spread = 5, SNR = 50 dB.

9. Conclusions

A new adaptive algorithm, called the SRLMF algorithm, has been presented in this work. Expressions are derived for the steady-state EMSE in a stationary environment. A condition for the mean convergence is also found, and it turns out that the convergence of the SRLMF algorithm strongly depends on the choice of initial conditions. Also, expressions are obtained for the tracking EMSE in a nonstationary environment. An optimum value of the step-size μ is also evaluated. Moreover, an extension of the weighted variance relation is provided in order to derive expressions for the mean-square error (MSE) and the mean-square deviation (MSD) of the proposed algorithm during the transient phase. Monte Carlo simulations have shown that there is a good agreement between the theoretical and simulated results. The simulation results indicate that both the SRLMF algorithm and the LMF algorithm converge at the same rate resulting in no performance loss. The analysis developed in this paper is believed to make practical contributions to the design of adaptive filters using the SRLMF algorithm instead of the LMF algorithm in pursuit of the reduction in computational cost and complexity whilst still maintaining good performance.

Acknowledgment

References

[1] H. Sari, "Performance evaluation of three adaptive equalization algorithms," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '82)*, vol. 7, pp. 1385–1389, May 1982.

[2] N. J. Bershad, "On the optimum data nonlinearity in LMS adaptation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 1, pp. 69–76, 1986.

[3] C. P. Kwong, "Dual sign algorithm for adaptive filtering," *IEEE Transactions on Communications*, vol. 34, no. 12, pp. 1272–1275, 1986.

[4] O. Macchi, "Advances in adaptive filtering," in *Digital Communications*, E. Biglieri and G. Prati, Eds., pp. 41–57, North-Holland, Amsterdam, The Netherlands, 1986.

[5] V. J. Mathews and S. H. Cho, "Improved convergence analysis of stochastic gradient adaptive filters using the sign algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 4, pp. 450–454, 1987.

[6] N. A. M. Verhoeckx and T. A. C. M. Claasen, "Some considerations on the design of adaptive digital filters equipped with the sign algorithm," *IEEE Transactions on Communications*, vol. 32, no. 3, pp. 258–266, 1984.

[7] E. Eweda, "Almost sure convergence of a decreasing gain sign algorithm for adaptive filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1669–1671, 1988.

[8] E. Eweda, "Tight upper bound of the average absolute error in a constant step-size sign algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 11, pp. 1774–1776, 1989.

[9] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "Comparison of the convergence of two algorithms for adaptive FIR digital filters," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 29, no. 3, pp. 670–678, 1981.

[10] N. J. Bershad, "Comments on 'comparison of the convergence of two algorithms for adaptive FIR digital filters'," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 6, pp. 1604–1606, 1985.

[11] W. A. Sethares, I. M. Y. Mareels, B. D. O. Anderson, C. R. Johnson Jr., and R. R. Bitmead, "Excitation conditions for signed regressor least mean squares adaptation," *IEEE Transactions on Circuits and Systems*, vol. 35, no. 6, pp. 613–624, 1988.

[12] E. Eweda, "Analysis and design of a signed regressor LMS algorithm for stationary and nonstationary adaptive filtering with correlated Gaussian data," *IEEE Transactions on Circuits and Systems*, vol. 37, no. 11, pp. 1367–1374, 1990.

[13] S. Dasgupta and C. R. Johnson Jr., "Some comments on the behavior of sign-sign adaptive identifiers," *Systems and Control Letters*, vol. 7, no. 2, pp. 75–82, 1986.

[14] S. I. Koike, "Analysis of the sign-sign algorithm based on Gaussian distributed tap weights," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '98)*, vol. 3, pp. 1673–1676, May 1998.

[15] D. L. Duttweiler, "Adaptive filter performance with nonlinearities in the correlation multiplier," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, no. 4, pp. 578–586, 1982.

[16] A. Gersho, "Adaptive filtering with binary reinforcement," *IEEE Transactions on Information Theory*, vol. 30, no. 2, pp. 191–199, 1984.

[17] S. Koike, "Effects of impulse noise at filter input on performance of adaptive filters using the LMS and signed regressor

LMS algorithms," in *Proceedings of the International Symposium on Intelligent Signal Processing and Communications (ISPACS '06)*, pp. 829–832, December 2006.

[18] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley Interscience, New York, NY, USA, 2003.

[19] E. Walach and B. Widrow, "The Least Mean Fourth (LMF) adaptive algorithm and its family," *IEEE Transactions on Information Theory*, vol. 30, no. 2, pp. 275–283, 1984.

[20] A. Zerguine, C. F. N. Cowan, and M. Bettayeb, "LMS-LMF adaptive scheme for echo cancellation," *Electronics Letters*, vol. 32, no. 19, pp. 1776–1778, 1996.

[21] T. Aboulnasr and A. Zerguine, "Variable weight mixed-norm LMS-LMF adaptive algorithm," in *Proceedings of the 33rd Annual Asilomar Conference on Signals, Systems, and Computers*, pp. 791–794, Pacific Grove, Calif, USA, October 1999.

[22] M. Moinuddin and A. Zerguine, "Tracking analysis of the NLMS algorithm in the presence of both random and cyclic nonstationarities," *IEEE Signal Processing Letters*, vol. 10, no. 9, pp. 256–258, 2003.

[23] A. Zerguine, M. K. Chan, T. Y. Al-Naffouri, M. Moinuddin, and C. F. N. Cowan, "Convergence and tracking analysis of a variable normalised LMF (XE-NLMF) algorithm," *Signal Processing*, vol. 89, no. 5, pp. 778–790, 2009.

[24] A. Zerguine, M. Moinuddin, and S. A. A. Imam, "A noise constrained least mean fourth (NCLMF) adaptive algorithm," *Signal Processing*, vol. 91, no. 1, pp. 136–149, 2011.

[25] R. Price, "A useful theorem for nonlinear devices having Gaussian inputs," *IRE Transactions on Information Theory*, vol. 4, no. 2, pp. 69–72, 1958.

[26] S. H. Cho, S. D. Kim, and K. Y. Jeon, "Statistical convergence of the adaptive least mean fourth algorithm," in *Proceedings of the 3rd International Conference on Signal Processing (ICSP '96)*, vol. 1, pp. 610–613, October 1996.