Proposed Hybrid Sparse Adaptive Algorithms for System Identification

Mahmood A. K Abdulsattar* Samer Hussein Ali**
*,**Department of Electrical Engineering /College of Engineering/ University of Baghdad
*Email:mahmood.abdulsattar@googlemail.com
**Email: msc_eng_samerhussein@yahoo.com

(Received 29 August 2016; accepted 19 December 2016)
https://doi.org/10.22153/kej.2017.12.003

Abstract

For sparse system identification, recent suggested algorithms are \( \ell_0 \)-norm Least Mean Square (\( \ell_0 \)-LMS), Zero-Attracting LMS (ZA-LMS), Reweighted Zero-Attracting LMS (RZA-LMS), and p-norm LMS (p-LMS) algorithms, that have modified the cost function of the conventional LMS algorithm by adding a constraint of coefficients sparsity. And so, the proposed algorithms are named \( \ell_0 \)-ZA-LMS, \( \ell_0 \)-RZA-LMS, p-ZA-LMS and p-RZA-LMS that are designed by merging two constraints from previous algorithms to improve the convergence rate and steady state of MSD for sparse system. In this paper, a complete analysis was done for the theoretical operation of proposed algorithms by exited white Gaussian sequence for input signal. The discussion of mean square deviation (MSD) with regard to parameters of algorithms and system sparsity was observed. In addition, in this paper, the correlation between proposed algorithms and the last recent algorithms were presented and the necessary conditions of these proposed algorithms were planned to improve convergence rate. Finally, the results of simulations are compared with theoretical study (?), which is presented to match closely by a wide-range of parameters.

Keywords: Adaptive filter, \( \ell_0 \)-LMS, zero-attracting, p-LMS, mean square deviation, Sparse system identification.

1. Introduction

Adaptive filtering is a difficult problem when the involved impulse response is sparse. Least Mean Square (LMS) algorithm used in many applications for example system identification, echo cancelation, and channel equalization, due to its easy implementation, good performance, and high robustness, [1–3].

The impulse response of a sparse system contains many zeros or near-zero coefficients and few large ones [4]. For this system, the conventional LMS not ever takes improvement of the sparsity. In latest years, a number of algorithms suggested depend on LMS to improve the feature of sparsity. The proportionate NLMS(PNLMS) algorithm and its improved ones, which use the individual step size with respect the coefficients of filter in proportional to improve the convergence rate of them[5, 6]. The recently developed sparse signal processing part [7–11]. The estimation of these algorithms is applied the features of unknown impulse response then added sparsity constraints to the cost function of gradient descent.

It is essential to deportment a theoretical examination of proposed algorithms. Numerical simulations proved the proposed algorithms have better performance than recent algorithms for sparse system identification, due to both minimized steady-state MSD and improved convergence rate.

1.1. Relation to Other Works

The mean square evaluation has been illustrated for LMS algorithm and a share of its deviations, the optimal algorithm parameters has
been selected to characterize the theoretically performance. LMS algorithm was proposed by Widrow for the first time in [12] and its performance was studied in [13]. Later, the mathematical frame of mean square analysis was established by Horowitz and Senne via examining the coefficients vector and succeeded the closed-form appearance of MSD [14], which was make simpler by Weinstein and Feuer [15]. A brief analysis, which planned in [16] depend on a style of adaptive algorithms, that presents linear time-invariant processes depend on the instantaneous of gradient vector and the LMS is the simplest algorithm. Likewise, the examination of Normalized LMS (NLMS) has involved extremely attention [17,18]. Even so, the styles stated, which are professional in their own background, might no longer be straight used in $\ell_0$-LMS, since its nonlinearity. The MSD analysis of ZA-LMS and p-LMS have been operated in [19, 20].

### 1.2. Main Contribution

The contribution which presented in this paper is on the performance analysis of steady-state. Then, the stability condition on step size is chosen. After that, the rule to select the parameter for steady-state performance is suggested. Finally, the steady-state MSD is achieved with the optimal parameter for proposed algorithms over the traditional algorithm.

Another contribution is on behavior of instantaneous analysis that implies the convergence rate for LMS algorithm. Also, the convergence rates of proposed algorithms are compared with that of standard LMS.

### 2. Background

#### 2.1. Standard LMS Algorithm

Let $h = [h_0 h_1 h_2 \ldots h_{L-1}]^T$ denotes the coefficient vector of filter, e.g., an impulse response of FIR filter; $x(n) = [x(n) x(n-1) \ldots x(n-L+1)]^T$ denotes a vector of input signal where $L$ is the filter length, $d(n)$ denotes a desired signal, $v(n)$ illustrates the observation noise vector, $y(n)$ denotes the output signal, $h_a(n)$ illustrates the estimated coefficient vector of adaptive filter at iteration $n$ and $e(n)$ denotes an estimation error.

$$d(n) = x^T(n) h + v(n) \quad \ldots(1)$$

$$y(n) = x^T(n) h_a(n) \quad \ldots(2)$$

$$e(n) = d(n) - y(n) \quad \ldots(3)$$

The cost function of traditional LMS $C(n)$ is expressed as

$$C(n) = \frac{1}{2} e(n)^2 \quad \ldots(4)$$

The filter coefficient vector for LMS is updated by:

$$h_a(n+1) = h_a(n) - \mu \frac{\partial C(n)}{\partial h_a(n)} \quad \ldots(5)$$

Where $\mu$ is the step size of adaptation [1].

### 3. $\ell_0$-LMS Algorithm

This algorithm was derived by inserting an $\ell_0$-norm constraint as a sparsity constraint to the cost function of traditional LMS to improve the convergence of LMS algorithm for sparse system identification [4].

The cost function is presented by the factorization

$$C(n) = \frac{1}{2} e(n)^2 + \gamma \epsilon_0 \sum_{i=1}^{L} (1 - e^{-\beta|h_{a_i}(n)|}) \quad \ldots(6)$$

The update of adaptive filter coefficient:

$$h_a(n+1) = h_a(n) + \mu e(n) x(n) - \rho \epsilon_0 s(n) \quad \ldots(7)$$

$$s(n) = \beta \text{sgn}(h_a(n)) e^{-\beta|h_a(n)|} \quad \ldots(8)$$

Where $\rho \epsilon_0 = \mu \gamma \epsilon_0$ is a parameter used to stabilize the estimation of error and the latest constraint, parameter $\beta$ is a positive value that is applied to define the area of zero attraction [4], $\text{sgn}(\cdot)$ is a component-wise sign function which defined as

---

Fig. 1. Adaptive System Identification (Direct System Modeling).

$$d(n) = x^T(n) h + v(n) \quad \ldots(1)$$

$$y(n) = x^T(n) h_a(n) \quad \ldots(2)$$

$$e(n) = d(n) - y(n) \quad \ldots(3)$$

The cost function of traditional LMS $C(n)$ is expressed as

$$C(n) = \frac{1}{2} e(n)^2 \quad \ldots(4)$$

The filter coefficient vector for LMS is updated by:

$$h_a(n+1) = h_a(n) - \mu \frac{\partial C(n)}{\partial h_a(n)} \quad \ldots(5)$$

Where $\mu$ is the step size of adaptation [1].

---

PDF created with pdfFactory trial version www.pdffactory.com
s_{gn}(x) = \begin{cases} 
\frac{x}{|x|} & x \neq 0 \\
0 & x = 0 
\end{cases} \quad \ldots(9)

By using the first order Taylor series of exponential functions, the computational complexity of (\(^r\)) can be reduced as followed \([4]\),
\[ e^{-\beta|x|} = \begin{cases} 
1 - \beta|x| & |x| \leq 1/\beta \\
0 & \text{elsewhere} 
\end{cases} \quad \ldots(10) 

Substituting (9) and (10) into (8), function \(s(\cdot)\) can be expressed clearly as
\[ s(n) = \begin{cases} 
-\beta^2 h_a(n) - \beta, & -1/\beta \leq h_{a1}(n) < 0 \\
-\beta^2 h_a(n) + \beta, & 0 \leq h_{a1}(n) < 1/\beta \\
0, & \text{elsewhere} \ldots(11) 
\end{cases} 

Fig. 2 describes the characteristic of the function \(s(x)\), which has zero-attraction effect.

![Fig. 2. Function \(s(n)\) for zero-attraction effect.](image)

4. (ZA-LMS) and (RZA-LMS) Algorithms

In the zero attractor, a cost function \(C_{ZA}(n)\) is defined by combining square error and \(l_1\) norm penalty of estimation coefficients vector as a sparsity constraint
\[ C_{ZA}(n) = \frac{1}{2}e(n)^2 + \gamma_{ZA}||h_a(n)||_1 \quad \ldots(12) 

The update filter of (ZA-LMS) is determined equally
\[ h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_{ZA} s_{gn}(h_a(n)) \quad \ldots(13) 

Where \( \rho_{ZA} = \mu \gamma_{ZA} \) is the factor used to control the force of sparsity penalty. In ZA-LMS all taps are forced to zero uniformly, and its performance will weaken in not sparse systems, then RZA-LMS use individual zero attractors for different filter taps and its cost function is
\[ C_{RZA}(n) = \frac{1}{2}e(n)^2 + \gamma_{RZA} \sum_{i=1}^{L} \log_{10}(1 + \epsilon|h_{a1}(n)|) \quad \ldots(14) 

The update filter of (RZA-LMS) is defined as
\[ h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_{RZA} s_{gn}(h_a(n)) \quad \ldots(15) 

Where \( \rho_{RZA} = \mu \gamma_{RZA} \epsilon \), where parameter \( \epsilon \) controls the similarity between (15) and (7).

5. p-LMS Algorithm

The cost function of (p-LMS) \(C_p(n)\) is defined by combining square error and \(l_p\) norm penalty the coefficient vector as shown as
\[ C_p(n) = \frac{1}{2}e(n)^2 + \gamma_p ||h_a(n)||_p \quad \ldots(16) 

The update equation of (P-LMS) as
\[ h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_p \frac{s_{gn}(h_a(n))}{1+|h_a(n)|^{1-p}} \quad \ldots(17) 

A parameter \( p \) has effect on the estimation bias in addition to the strength of sparsity correction. A parameter \( \rho_p = \mu \gamma_p \) is used to stabilize the constraint term and the estimation square error.

6. Proposed Algorithms

A. \(\ell_0\)-ZA-LMS and \(\ell_0\)-RZA-LMS Algorithms

The proposed cost functions of \(\ell_0\)-ZA-LMS and \(\ell_0\)-RZA-LMS are designed by merging between \(\ell_0\)-norm with \(\ell_1\)-norm on the coefficients of filter constraints and/or reweighted zero attractor constraints into the cost function of LMS to improve the convergence rate of it. They obviously present in (18) and (19).
\[ C(n) = \frac{1}{2}e(n)^2 + \gamma_{ZA} ||h_a(n)||_1 + \gamma_{LA} \sum_{i=1}^{L} (1 - \epsilon|h_{a1}(n)|) \quad \ldots(18) 

\[ C(n) = \frac{1}{2}e(n)^2 + \gamma_{RZA} \sum_{i=1}^{L} \log_{10}(1 + \epsilon|h_{a1}(n)|) + \gamma_{LA} \sum_{i=1}^{L} (1 - \epsilon|h_{a1}(n)|) \quad \ldots(19) 

The update of adaptive filter coefficient
\[ h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_{ZA} s_{gn}(h_a(n)) - \mu \rho_{ZA} \epsilon \frac{s_{gn}(h_a(n))}{1+|h_a(n)|} \quad \ldots(20) 

\[ h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_{RZA} \frac{s_{gn}(h_a(n))}{1+|h_a(n)|} \quad \ldots(21) 

B. p-ZA-LMS and p-RZA-LMS Algorithms

This proposed algorithms have designed by merging between \(p\)-norm with \(\ell_1\)-norm constraints on the coefficients of filter and/or
reweighted zero attractor constraints into the LMS cost function to increase the convergence of LMS for sparse systems. They obviously present in (22) and (23).

\[
C(n) = \frac{1}{2} e(n)^2 + \gamma_{ZA} ||h_a(n)||_1 + \gamma_p ||h_a(n)||_p^p
\]  

\[
C(n) = \frac{1}{2} e(n)^2 + \gamma_{ZA} \sum_{i=1}^{L} \log_{10}(1 + e|h_{a_i}(n)|) + \gamma_p ||h_a(n)||_p^p
\]  

(22)

(23)

The update of adaptive filter coefficient

\[
h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_{ZA} sgn(h_a(n)) - \rho_p \frac{sgn(h_a(n))}{1+|h_a(n)|^{1-p}}
\]  

(24)

\[
h_a(n+1) = h_a(n) + \mu e(n)x(n) - \rho_{ZA} sgn(h_a(n)) + \rho_p \frac{sgn(h_a(n))}{1+|h_a(n)|^{1-p}}
\]  

(25)

7. Simulation Results

We illustrated the performance of the proposed algorithms for system identification via a computer simulation. An impulse response of unknown system consists 16 coefficients may be one of three systems, the first impulse response, the value of 5th tap equal 1 and the others equal zero is called sparse system, the second impulse response, the values of odd taps equal 1 and the others equal zero is called semi sparse system and the third impulse response, the values of all taps equal 1 is called not sparse system. A white Gaussian noise used as input signal and observed noise with variances 1 and 0.01 individually.

The first experiment is planned to examine the convergence rate performance of sparse system with apply our methods with different value of \(\mu\). The parameters of algorithms are providing in Tables 1, 2, 3, and 4. The results of algorithms are achieved from independent simulations, as shown in Fig’s.3, 4, .5, and .6, these are obvious that the convergence rate of proposed algorithms are more rapidly and produces lower MSD than the LMS are done in a large value of \(\rho\).

The second experiment is planned to test the performance of the proposed algorithms via various sparsity. The unknown system here is sparse system, then after 1500 iterations are semi sparse system and later, after 3000 iteration be not sparse system. The parameters are set as in table 5. Fig. 7 and Fig. 8 show the average estimate of mean square deviation (MSD). Both the \(\ell_0\)-RZA-LMS and the p-RZA-LMS return better steady-state MSD and faster convergence than other algorithms (before the 1500th iteration) when the system is sparse. When the number of non-zero taps increases to 8, (after the 1500th iteration and before the 3000th iteration) when the system is semi sparse, the performance of algorithms deteriorate while the \(\ell_0\)-RZA-LMS and p-RZA-LMS maintains the best performance and the MSD of \(\ell_0\)-RZA-LMS algorithm is lower than that of \(\ell_0\)-RZA-LMS algorithm and the MSD of p-RZA-LMS algorithm is lower than that of p-ZA-LMS algorithm among this filter. When the system is non-sparse (after 3000 iterations), the \(\ell_0\)-ZA-LMS and p-RZA-LMS maintain the best performance with this filter while the others still achieves comparably to the LMS.

The third experiment suggests a system with 128-taps with 8 nonzero coefficients as shown in Fig 9. The iterations of all filters are 5000. Table 6 present the parameters of algorithms for this experiment, the average MSD is shown in Fig.10 and Fig.11. For this long sparse system, the convergence rate of all filters is almost the best form that of LMS, but the MSD of \(\ell_0\)-RZA-LMS and p-RZA-LMS are relatively minimum.

<table>
<thead>
<tr>
<th>Table 1, Parameters of (\ell_0)-ZA-LMS Algorithm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
</tr>
<tr>
<td>LMS</td>
</tr>
<tr>
<td>(\ell_0)-ZA-LMS2</td>
</tr>
<tr>
<td>(\ell_0)-ZA-LMS3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2, Parameters of (\ell_0)-RZA-LMS Algorithm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
</tr>
<tr>
<td>LMS</td>
</tr>
<tr>
<td>(\ell_0)-RZA- LMS2</td>
</tr>
<tr>
<td>(\ell_0)-RZA- LMS3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3, Parameters of p-ZA-LMS Algorithm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
</tr>
<tr>
<td>LMS</td>
</tr>
<tr>
<td>p-ZA-LMS2</td>
</tr>
<tr>
<td>p-ZA-LMS3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4, Parameters of p-RZA-LMS Algorithm.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
</tr>
<tr>
<td>LMS</td>
</tr>
<tr>
<td>p-RZA-LMS2</td>
</tr>
<tr>
<td>p-RZA-LMS3</td>
</tr>
</tbody>
</table>
Table 5, Parameters of Algorithms for Second Experiment.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>μ</th>
<th>ρ_ε₀</th>
<th>β</th>
<th>p_{ZA}</th>
<th>p_{RZA}</th>
<th>ε</th>
<th>ρ_p</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMS</td>
<td>0.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ₀-LMS</td>
<td>0.025</td>
<td>8 × 10^{-5}</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZA-LMS</td>
<td>0.025</td>
<td>1.6 × 10^{-4}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RZA-LMS</td>
<td>0.025</td>
<td></td>
<td>2.5 × 10^{-4}</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-LMS</td>
<td>0.025</td>
<td></td>
<td></td>
<td></td>
<td>4 × 10^{-4}</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ₀-ZA-LMS</td>
<td>0.025</td>
<td>2.5 × 10^{-4}</td>
<td>5</td>
<td>2.5 × 10^{-5}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ₀-RZA-LMS</td>
<td>0.025</td>
<td>1.3 × 10^{-4}</td>
<td>5</td>
<td>2.5 × 10^{-4}</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-ZA-LMS</td>
<td>0.025</td>
<td></td>
<td>2.5 × 10^{-5}</td>
<td></td>
<td>2.5 × 10^{-4}</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-RZA-LMS</td>
<td>0.025</td>
<td></td>
<td>2.5 × 10^{-5}</td>
<td></td>
<td>2.5 × 10^{-4}</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6, Parameters of Algorithms for Third Experiment.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>μ</th>
<th>ρ_ε₀</th>
<th>β</th>
<th>p_{ZA}</th>
<th>p_{RZA}</th>
<th>ε</th>
<th>ρ_p</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMS</td>
<td>0.0078</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ₀-LMS</td>
<td>0.0078</td>
<td>2.5 × 10^{-6}</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZA-LMS</td>
<td>0.0078</td>
<td>6 × 10^{-5}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RZA-LMS</td>
<td>0.0078</td>
<td></td>
<td>2.5 × 10^{-5}</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-LMS</td>
<td>0.0078</td>
<td></td>
<td></td>
<td></td>
<td>8 × 10^{-5}</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ₀-ZA-LMS</td>
<td>0.0078</td>
<td>8 × 10^{-6}</td>
<td>5</td>
<td>8 × 10^{-6}</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ℓ₀-RZA-LMS</td>
<td>0.0078</td>
<td>8 × 10^{-6}</td>
<td>5</td>
<td>2 × 10^{-5}</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-ZA-LMS</td>
<td>0.0078</td>
<td></td>
<td>8 × 10^{-7}</td>
<td></td>
<td>8 × 10^{-5}</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-RZA-LMS</td>
<td>0.0078</td>
<td></td>
<td>6 × 10^{-5}</td>
<td></td>
<td>2.5 × 10^{-6}</td>
<td>0.6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. Learning curves of ℓ₀-ZA-LMS with different ρ, driven by white signal.

Fig. 4. Learning curves of ℓ₀-RZA-LMS with different ρ, driven by white signal.

Fig. 5. Learning curves of p-ZA-LMS with different ρ, driven by white signal.

Fig. 6. Learning curves of p-RZA-LMS with different ρ, driven by white signal.
Fig. 7. The performance of different algorithms of varying sparsity, driven by white signal.

Fig. 8. The performance of different algorithms of varying sparsity, driven by white signal.

Fig. 9. 128-order adaptive filter.

Fig. 10. The performance of 128-order adaptive filters, driven by white input signal.

Fig. 11. The performance of 128-order adaptive filters, driven by white input signal.

8. References


الخوارزميات التكيفية الهجينة المقترحة لتحديد هوية الأنظمة التكيفية المتناثرة

محمود عبد القادر عبد الستار* سامر حسين علي**

قسم الهندسة الكهربائية، كلية الهندسة، جامعة بغداد

البريد الإلكتروني: mahmood.abdulsattar@gmail.com

البريد الإلكتروني: msc_eng_samerhussein@yahoo.com

الخلاصة

لتحديد النظام التكيفي المتناثر اقترحنا سابقاً خوارزميات وتمثلت في 

- وصف الخوارزميات المقترحة

-

- ووصف الخوارزميات المقترحة

وأُسمى الخوارزميات المقترحة حالياً في LMS، وذلك بالإضافة إلى مناسب لدالة تدفق LMS، ونستخدم الخوارزميات المقترحة حالياً في LMS ووصف الخوارزميات المقترحة.

وقد اقترحنا استخدام الخوارزميات المقترحة لتحديد هوية الأنظمة التكيفية المتناثرة. وفي هذا البحث يوجد تحديداً ميكانيك للخوارزميات المقترحة وحل الخوارزميات المقترحة النسبية التناظرية للخوارزميات المتنوعة حالياً، والتي تبين

تحسين معدل التقارب للخوارزميات المقترحة. وأخيراً، هناك مقارنة بين نتائج اقتصاد الأنظمة المحاكاة والنتائج النظرية والاعتماد على المعلومات لكل خوارزمية.