Gradient Comparator Least Mean Square Algorithm for Identification of Sparse Systems with Variable Sparsity

Bijit Kumar Das[1] and Mrityunjoy Chakraborty [2]
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur, INDIA
Phone: +91-3222-283512 Fax: +91-3222-255303
E.Mail : [1]bijitbijit@gmail.com, [2] mrityun@ece.iitkgp.ernet.in

***Abstract*—In adaptive system identification, exploitation of sparsity that may be inherent in the system leads to improved performance of the identification algorithms. The recently proposed ZA-LMS algorithm achieves this by introducing a "zero attractor" term in the update equation that tries to pull the coefficients towards zero, thus accelerating the convergence. For systems whose sparsity level, however, varies over a wide range, from highly sparse to non-sparse, the ZA-LMS algorithm, however, performs poorly, as it can not distinguish between the zero and the non-zero taps of the system. In this paper, we propose a modified ZA-LMS algorithm for tackling the case of variable sparseness, which selectively chooses the zero attractors only for the "inactive" taps. The proposed method is very simple, easy to implement and well supported by simulation studies.**

I. INTRODUCTION

In practice, one often encounters systems that have a sparse impulse response, with the degree of sparseness varying with time and context. Examples of such systems include echo paths for both network and acoustic echo [1]-[2], sparse wireless multipath channels [3] and shallow underwater channels [4] for acoustic communication. Conventional adaptive system identification algorithms like the LMS [5] and its variants, however, do not make use of the a priori knowledge of the sparseness of the system and thus perform poorly both in terms of steady state excess mean square error (EMSE) and convergence speed. In recent years, several algorithms have been proposed that exploit the sparsity of the system and achieve better performance, like the partial update LMS deploying either statistical detection of active taps [6]-[7] or sequential partial updating [8]-[9], the proportionate normalized LMS (PNLMS) and its variants [10]-[11] etc. More recently, motivated by LASSO [12] and the recent progresses in compressive sensing [13]-[14], an alternative approach has been proposed in [15] to identify sparse systems which introduces a l_1 norm (of the coefficients) penalty in the cost function which favors sparsity. This results in a modified LMS update with a zero attractor for all the taps, named as the Zero-Attracting LMS (ZA-LMS). A variant of the ZA-LMS is also proposed in [15] which employs reweighted step sizes for the different taps to adjust to variable sparsity. This is, however, associated with huge computational burden due to the L division operations at each step, where L is the number of taps in the filter. Separately, in [16], a class of sparseness-controlled algorithms have been proposed, which are robust against the variation of sparsity in the system model while requiring again much higher computational complexity.

In this paper, we propose a modified ZA-LMS algorithm for tackling the case of variable sparseness. The proposed algorithm uses a gradient comparison mechanism for comparing the sign of the standard LMS part in the ZA-LMS update terms, with the sign of the so-called "zero attractor" and selectively chooses the zero attractors only for the "inactive" taps. The proposed method is very simple and easy to implement, with simulation studies confirming its improved performance over the conventional ZA-LMS algorithm.

II. REVIEW OF THE ZA-LMS ALGORITHM

In a ZA-LMS based adaptive filter, with $x(n)$ as input and $d(n)$ as the desired response, a L-th order filter coefficient vector $\mathbf{w}(n) = [w_0(n), w_1(n), \cdots, w_{L-1}(n)]^T$ is updated in time as [15],

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) - \rho sgn[\mathbf{w}(n)], \quad (1)$$

where $\mathbf{x}(n) = [x(n), x(n-1), \cdots, x(n-L+1)]^T$ is the input vector, $y(n) = \mathbf{w}^T(n)\mathbf{x}(n)$ is the filter output, $e(n) = d(n) - y(n)$ is the filter output error, all at the n^{th} time index, ρ is a suitably chosen constant with very small magnitude and μ is the usual adaptation step size. In a system identification problem, $d(n)$ is the signal observed at the plant output, i.e., $d(n) = \mathbf{w}_{opt}^T\mathbf{x}(n)+e_{opt}(n)$, where $\mathbf{w}_{opt}$ is the true system coefficient vector and $e_{opt}(n)$ is the observation noise, independent of $x(n)$ and with variance $\sigma_v^2 = E[e_{opt}^2(n)]$. Like the LMS algorithm, the ZA-LMS too converges in mean though with a bias, under identical convergence condition, namely, $0 < \mu < \frac{2}{Tr(\mathbf{R})}$, where $\mathbf{R} = E[\mathbf{x}(n)\mathbf{x}^T(n)]$ is the input autocorrelation matrix. The limiting value of the adaptive filter coefficient vector is given by [15],

$$\lim_{n\to\infty} \mathbf{w}(n) = \mathbf{w}_{opt} - \frac{\rho}{\mu}\mathbf{R}^{-1}E[sgn[\mathbf{w}(n)]], \quad (2)$$

where $sgn[.]$ is the usual signum function. The term $-\frac{\rho}{\mu}\mathbf{R}^{-1}E[sgn[\mathbf{w}(n)]]$ in the R.H.S. of (2) provides the bias term. As compared to the standard LMS algorithm, the ZA-LMS update equation (1) has an extra update term, $-\rho sgn[\mathbf{w}(n)]$. This is the so-called zero attractor which tries to pull the filter coefficients to zero and as a result, for the zero taps, the convergence is accelerated. For the non-zero taps, however, while the standard LMS update term in (1), namely, $\mu e(n)\mathbf{x}(n)$ aims at convergence (in mean) of the tap weights to their optimum values, the zero-attractor still tries to attract the coefficients to zero, leading to slowing down of the convergence for these taps. This is one major drawback of the ZA-LMS algorithm, as it can not distinguish between sparse and non-sparse systems. In this paper, we propose a novel gradient comparison based mechanism to improve the convergence properties of the ZA-LMS algorithm, making it robust against variable system sparsity.

III. The proposed Gradient Comparison Based Algorithm

In the proposed gradient comparison based LMS (GC-LMS) algorithm, the zero-attractors are selectively chosen for only those taps that have polarity same as that of the gradient of the squared instantaneous error. The compact form of the proposed GC-LMS weight update equation is given by,

$$\begin{aligned}&\mathbf{w}(n+1)\\ &\quad = \mathbf{w}(n) + \mu e(n)\mathbf{x}(n) - \rho \mathbf{D}(n) sgn(\mathbf{w}(n)),\end{aligned} \tag{3}$$

where $\mathbf{D}(n)$ is a diagonal matrix with $|0.5\{sgn(e(n)\mathbf{x}(n)) - sgn(\mathbf{w}(n))\}|$ vector as diagonal, i.e., $\mathbf{D}(n) = diag[|0.5\{sgn(e(n)\mathbf{x}(n)) - sgn(\mathbf{w}(n))\}|]$.

To understand the functioning of the algorithm consider the diagonal of $\mathbf{D}(n)$. The first term in this is given by $sgn(e(n)\mathbf{x}(n))$ where the vector $e(n)\mathbf{x}(n)$ provides the negative of the gradient of the squared instantaneous error, $e^2(n)$ with respect to the tap weights at $\mathbf{w} = \mathbf{w}(n)$. More specifically, $e(n)\mathbf{x}(n) = -\frac{1}{2}[\frac{\partial e^2(n)}{\partial w_0} \cdots \frac{\partial e^2(n)}{\partial w_{L-1}}]^T|_{\mathbf{w}=\mathbf{w}(n)}$. Now, consider the i^{th} tap weight $w_i(n)$, $i = 0, 1, \cdots, L-1$. Let the corresponding true system coefficient be denoted by $w_{opt,i}$. Then, first consider the case where $w_{opt,i} \neq 0$. Of the two possibilities, $w_{opt,i} > 0$ and $w_{opt,i} < 0$, we consider one of them, say, $w_{opt,i} > 0$ (the same logic would apply to the case $w_{opt,i} < 0$ as well). If, suppose $w_i(n) > w_{opt,i}$, then, on an average, from the quadratic nature of the $e^2(n)$ surface (as a function of $w_0(n)$, $w_1(n)$,...,$w_{L-1}(n)$), $\frac{\partial e^2(n)}{\partial w_i(n)}$ will be positive, i.e., $[e(n)\mathbf{x}(n)]_i$ will be negative, meaning $[\mathbf{D}(n)]_{i,i} = 1$, and thus, for the $i^{th} tap$, both the zero-attractor term $-\rho[sgn(\mathbf{w}(n))]_i$ and the LMS update term $\mu[e(n)\mathbf{x}(n)]_i$ will be acting in the same sense. If, on the other hand, $w_i(n) < w_{opt,i}$, on an average, $\frac{\partial e^2(n)}{\partial w_i(n)}$ will be negative, i.e., $[e(n)\mathbf{x}(n)]_i$ will be positive. Now, if $w_i(n)$ is also positive, we will have $[\mathbf{D}(n)]_{i,i} = 0$, meaning, the effect of the zero-attractor, $-\rho sgn[w_i(n)]$ to pull $w_i(n)$ towards zero *will be neutralized* and the update equation (3) will follow the standard LMS update dynamics. If, on the other hand, $w_i(n) < 0$, by the same logic as above, $[\mathbf{D}(n)]_{i,i} = 1$, meaning both the zero-attractor $-\rho sgn(w_i(n))$ and the LMS update term, $\mu[e(n)\mathbf{x}(n)]_i$ will act in the same sense. If, however, $w_{opt,i} = 0$, we always have $[\mathbf{D}(n)_{i,i}] = 1$ and the update equation (3) turns out to be the simple ZA-LMS algorithm.

So, the main goal for using the diagonal $\mathbf{D}(n)$ is not to use the zero attractor whenever it is working against the conventional LMS update, because the use of zero attractor is only advantageous, in mean square error sense, if it strengthens the standard LMS update dynamics.

Like the LMS algorithm and its several variants, the proposed GC-LMS algorithm too converges in mean. However, like the ZA-LMS algorithm, it too converges with a bias, as explained in the theorem below.

Theorem 1 : In the proposed GC-LMS algorithm, the mean coefficient vector $E[\mathbf{w}(n)]$ converges as $n \to \infty$ for $0 < \mu < \frac{2}{Tr(\mathbf{R})}$, with

$$\begin{aligned}\lim_{n\to\infty} E[\mathbf{w}(n)] &= E[\mathbf{w}(\infty)]\\ &= \mathbf{w}_{opt} - \frac{\rho}{\mu}\mathbf{R}^{-1}E[\mathbf{D}(\infty)sgn[\mathbf{w}(\infty)]].\end{aligned} \tag{4}$$

Proof: Omitted.

IV. Simulation Studies

The proposed algorithm was simulated for identifying a FIR system with 16 tap coefficients. Initially, the system was taken to be highly sparse with only one non-zero element, -0.9 in the first tap position and all other taps set to zero. After 1000 time steps, the system was changed to a semi-sparse system that has non-zero values only for the first five tap coefficients, given as -0.9, 0.8,-0.5,0.6,-0.4. Finally, after 2000 time steps, the system was changed to a non-sparse system, with $\mathbf{w}_{opt} = [-0.9, 0.5, 0.5, 0.5, 0.5, 0.5, -0.76, -0.76, -0.76, -0.76, 0.3, 0.9, -0.4, 0.4, -0.6, 0.7]$. The simulation was carried out for a total of 3000 iterations, with $\mu = 0.07$, $\rho = 0.0015$, $\sigma_v^2 = 0.001$, and with the input $x(n)$ taken as a zero mean, unit variance white random process. The simulation results are shown in Fig. 1 by plotting the MSE against the iteration index n (obtained by averaging $e^2(n)$ over 400 experiments) for the proposed algorithm (green line) against that for the ZA-LMS (red line) and the standard LMS (blue line) algorithms. It is easily seen from Fig.(1) that for sparse and semi-sparse system models, the proposed algorithm performs at par with the better of the ZA-LMS and the LMS algorithms (former for sparse systems and latter for semi-sparse systems). For highly non-sparse systems, however the performance of the GC-LMS, though better than the ZA-LMS, turns out to be somewhat inferior to that of the standard LMS algorithm. Finally we have also simulated in Fig. 2 the variation of the steady-state EMSE (for a particular convergence speed) for the LMS (blue), the ZA-LMS (red) and the proposed GC-LMS (green) algorithms as functions of system sparsity (defined as the ratio of the number of nonzero taps to the number of total taps, which varies in the range of 0 to 1 and which actually decreases with increasing sparsity.)

V. CONCLUSIONS

This paper addresses the issue of identification of sparse systems with variable sparsity. For this, a new adaptive filter has been proposed by modifying the recently proposed ZA-LMS algorithm, which introduces certain "zero attractors", selectively for the unknown zero taps of the system, thus accelerating the overall convergence. The proposed method is very simple, easy to implement and well supported by simulation studies.

REFERENCES

[1] J. Radecki, Z. Zilic, and K. Radecka, "Echo cancellation in IP networks", *Proc. Fourty-Fifth Midwest Symposium on Circuits and Systems*, vol. 2, 2002, pp. 219-222.

[2] E. Hansler, "The hands-free telephone problem- an annotated bibliography", *Signal Processing*, vol. 27, no. 3, pp. 259-271, Jun. 1992.

[3] Ian J. Fevrier, Saul B. Gelfand, and Michael P. Fitz, "Reduced Complexity Decision Feedback Equalization for Multipath Channels with Large Delay Spreads", *IEEE Trans. Communications*, vol. 47, No. 6, pp. 927-937, June 1991.

[4] Subhadeep Roy, Tolga M. Duman, and Vincent Keyko McDonald, "Error Rate Improvement in Underwater MIMO Communications Using Sparse Partial Response Equalization ", *IEEE Journal of Oceanic Engineering*, vol. 34, No. 2, pp. 181-201, April 2009.

[5] S. Haykin, *Adaptive Filter Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1986.

[6] J. Homer, I. Mareels, R.R. Bitmead, B. Wahlberg, and A. Gustafsson, LMS estimation via structural detection, *IEEE Trans. on Signal Processing*, vol. 46, pp. 26512663, October 1998.

[7] Y. Li, Y. Gu, and K. Tang, Parallel NLMS filters with stochastic active taps and step-sizes for sparse system identification, in *Proceedings of ICASSP*, 2006, vol. 3, pp. 109112.

[8] D.M. Etter, Identification of sparse impulse response systems using an adaptive delay filter, in *Proceedings of ICASSP*, 1985, pp. 11691172.

[9] M. Godavarti and A. O. Hero, Partial update LMS algorithms, *IEEE Trans. on Signal Processing*, vol. 53, pp. 2382 2399, 2005.

[10] D. L. Duttweiler, Proportionate normalized least mean square adaptation in echo cancellers, *IEEE Trans. Speech Audio Processing*, vol. 8, no. 5, pp. 508518, Sep. 2000.

[11] S.L. Gay, An efficient, fast converging adaptive filter for network echocancellation, in *Proceedings of Asilomar*, 1998, vol. 1, pp. 394398.

[12] R. Tibshirani, "Regression shrinkage and selection via the lasso," in *J. Royal. Statist. Soc B.,* vol. 58, pp. 267-288, 1996.

[13] E. Candes, "Compressive Sampling" in *Int. Congress of Mathematics*, vol. 3, pp. 1433-1452, 2006.

[14] R. Baraniuk, "Compressive sensing," in *IEEE Signal Processing Magazine*, vol. 25, pp. 21-30, March 2007.

[15] Y. Gu Y. Chen and A. O. Hero, "Sparse LMS for system identification," in *Proc. IEEE Intl. Conf. Acoust. Sp. Sig. Proc., Taipei, Taiwan,* Apr. 2009.

[16] P. Loganathan, A. W. H. Khong, P. A. Naylor, "A class of sparseness-controlled algorithms for echo cancellation", *IEEE Trans. Audio, Speech and Language Processing*, Vol. 17, NO.8, pp. 1591-1601, Nov. 2009

[17] B. Farhang-Boroujeny, *Adaptive Filters*, John Wlley and Sons.

[18] A. H. Sayed, *Adaptive Filters*, John Wiley and Sons, 2008.

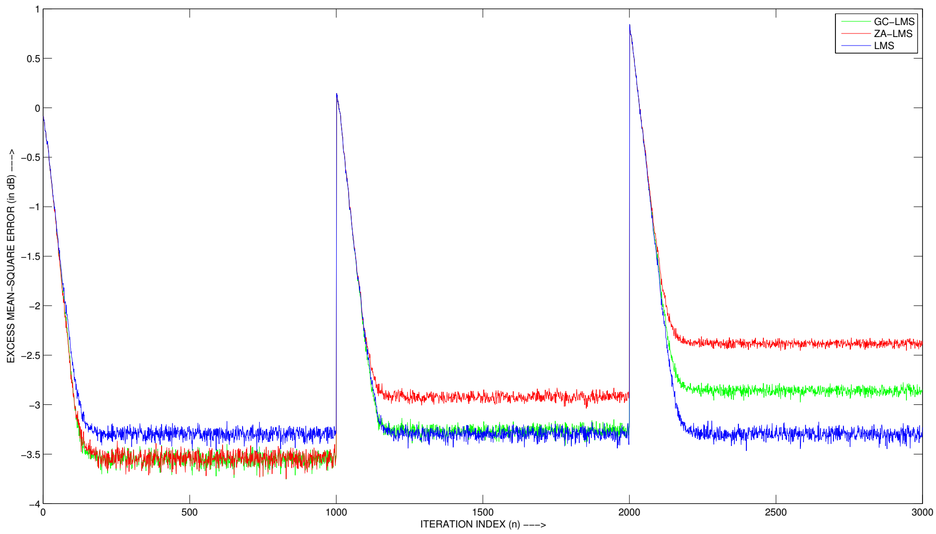

Fig. 1. The MSE versus no. of observations for the standard LMS (blue), the ZA-LMS (red) and the proposed GC-LMS (green) algorithms.

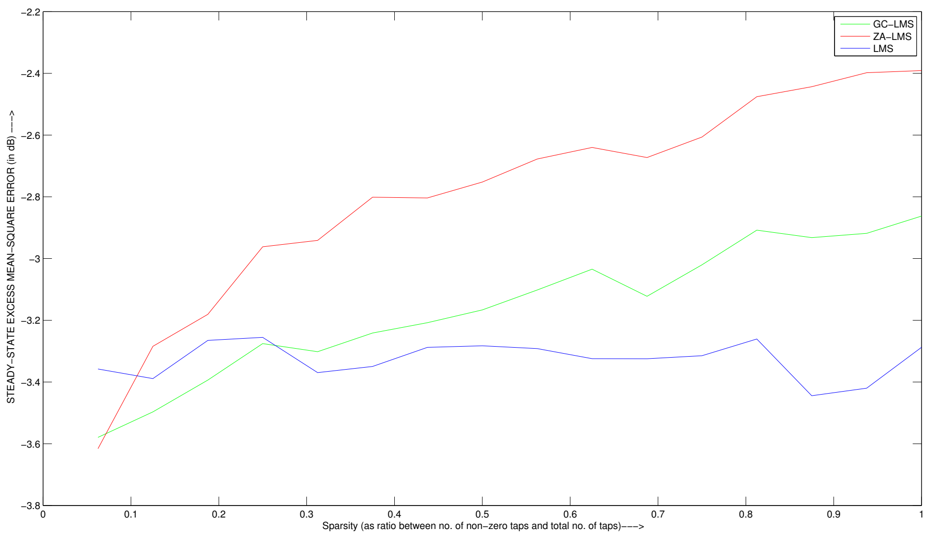

Fig. 2. The steady-state EMSE versus system sparsity for the standard LMS (blue), the ZA-LMS (red) and the proposed GC-LMS (green) algorithm.