LOW COST VARIABLE STEP-SIZE LMS WITH MAXIMUM SIMILARITY TO THE AFFINE PROJECTION ALGORITHM

Miguel ferrer, Maria de Diego, Alberto Gonzalez
agonzal@dcom.upv.es

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Abstract

The LMS algorithm is widely employed in adaptive systems due to its robustness, simplicity, and reasonable performance. However, it is well known that this algorithm suffers from a slow convergence speed when dealing with colored reference signals. The affine projection algorithm is a good alternative in this case. This algorithm has the peculiarity of starting from N data vectors of the reference signal. It transforms these vectors into as many data vectors suitably normalized in energy and mutually orthogonal. In this work, we propose a version of the LMS algorithm that, similar to the affine projection algorithm, starts from N data vectors of the reference signal but corrects them by using only a scalar factor that functions as a convergence step. Our goal is to align the behavior of this algorithm with the behavior of the affine projection algorithm without significantly increasing the computational cost of the LMS. [1]

Algorithms update equations

NLMS:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \frac{1}{\mathbf{x}_L^T(n)\mathbf{x}_L(n)}\mathbf{x}_L(n)e^a(n). \quad (1)$$

AP:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{X}(n)[\mathbf{X}^T(n)\mathbf{X}(n)]^{-1}\mathbf{e}^a(n), \quad (2)$$

$$\mathbf{X}(n) = [\mathbf{x}_L(n), \mathbf{x}_L(n-1), \cdots, \mathbf{x}_L(n-N+1)].$$

APL:

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mu(n)\mathbf{X}(n)\mathbf{e}^a(n), \quad (3)$$

with

$$\mu(n) = \mu_I(n) = \frac{\|\mathbf{X}(n)\mathbf{e}^a(n)\|^2}{\|\mathbf{X}^T(n)\mathbf{X}(n)\mathbf{e}^a(n)\|^2}, \quad (4)$$

defines the affine-like-I (APL-I) [2].

Conclusion

In this work, it is proposed a modification of the LMS algorithm that behaves like the AP avoiding matrix inversion. It is a robust algorithm that significantly improves the performance of NLMS with minimal additional computational cost.
It exhibits excellent performance for colored signals up to the projection orders where the convergence behavior of the AP cannot be improved either. Therefore, the algorithm's performance is significant, and the trade-off between convergence speed and computational cost is, in most cases, much superior to that of other similar algorithms such as NLMS or other APL proposals.

References

[1] Miguel Ferrer, María de Diego, and Alberto Gonzalez. Low cost variable step-size lms with maximum similarity to the affine projection algorithm. *IEEE Open Journal of Signal Processing*, 5:82–91, 2024.

[2] Md. Zulfiquar Ali Bhotto and Andreas Antoniou. Affine-projection-like adaptive-filtering algorithms using gradient-based step size. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 61(7):2048–2056, 2014.

Proposed variable step-size selection

It may be inferred that an adaptive algorithm would exhibit similar behaviour to a given one if its coefficients were very close at each algorithm iteration. Therefore, we propose the use of a variable convergence step that minimises the squared 2-norm of the difference between the coefficients of the exact AP algorithm, denoted as $\mathbf{w}_{AP}(n)$ and shown in (2), and the approximate version, denoted as $\mathbf{w}_{APL}(n)$ and shown in (3). This means

$$\tilde{\mu}(n) = \arg\min_{\mu(n)} \left\{\|\mathbf{w}_{AP}(n) - \mathbf{w}_{APL}(n)\|^2\right\}, \quad (5)$$

or equivalently

$$\tilde{\mu}(n) = \arg\min_{\mu(n)} \left\{\|\mathbf{X}(n)(\mathbf{X}(n)^T\mathbf{X}(n))^{-1} - \mu(n)\mathbf{X}(n))\mathbf{e}^a(n)\|^2\right\}, \quad (6)$$

leading to

$$\tilde{\mu}(n) = \frac{(\mathbf{e}^a(n))^T\mathbf{e}^a(n)}{[\mathbf{X}(n)\mathbf{e}^a(n)]^T\mathbf{X}(n)\mathbf{e}^a(n)} = \frac{\|\mathbf{e}^a(n)\|^2}{\|\mathbf{e}^a(n)\|^2_{\mathbf{\Sigma}(n)}}, \quad (7)$$

where $\mathbf{\Sigma}(n) = \mathbf{X}(n)^T\mathbf{X}(n)$. Thus, the proposed approach uses the update equation in (3), just like the AP and the APL-I, except that the convergence step is obtained by solving the minimization problem in (5). This approach would require $LN + 3N$ multiplications for updating the coefficients, which is a lower count compared to the AP and the APL-I algorithms.

Convergence discussion

It is suggested in [2] that:

$$0 < \mu(n) < \mu_{\max}(n) = \frac{2}{\lambda_{\max}(n)}, \quad (8)$$

where $\lambda_{\max}(n)$ is the maximum eigenvalue of $\mathbf{X}(n)\mathbf{X}^T(n)$ (equivalently $\lambda_{\min}(n)$).

Equation (7) is a generalised Rayleigh quotient. This ensures that the following boundaries are satisfied

$$\frac{1}{\lambda_{\max}(n)} \leq \tilde{\mu}(n)\,\{\mu_I(n)\} \leq \frac{1}{\lambda_{\min}(n)}. \quad (9)$$

We can only guarantee that (8) is fulfilled when the eigenvalues of the matrix $\mathbf{X}^T(n)\mathbf{X}(n)$ are not sparse.

$$\lambda_{\max}(n)/\lambda_{\min}(n) < 2, \quad (10)$$

which stands for low colored signals and colored ones and low projection orders.

The MSE can be approximated as

$$\mathrm{MSE} \approx \sigma_v^2\frac{3\mathrm{N}-1}{2\mathrm{N}-1}, \quad (11)$$

and MSE $< 2\sigma_v^2$.

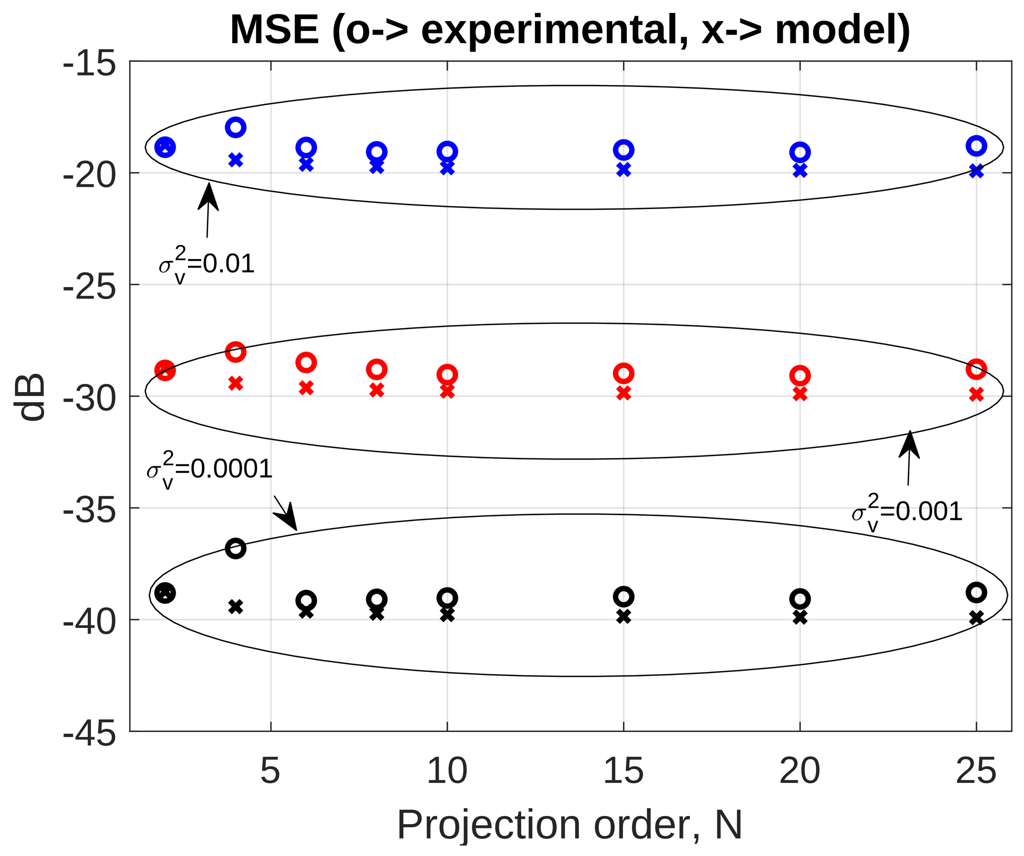

Figure 1: Experimental and theoretical MSE versus projection order for the system identification problem.

Results

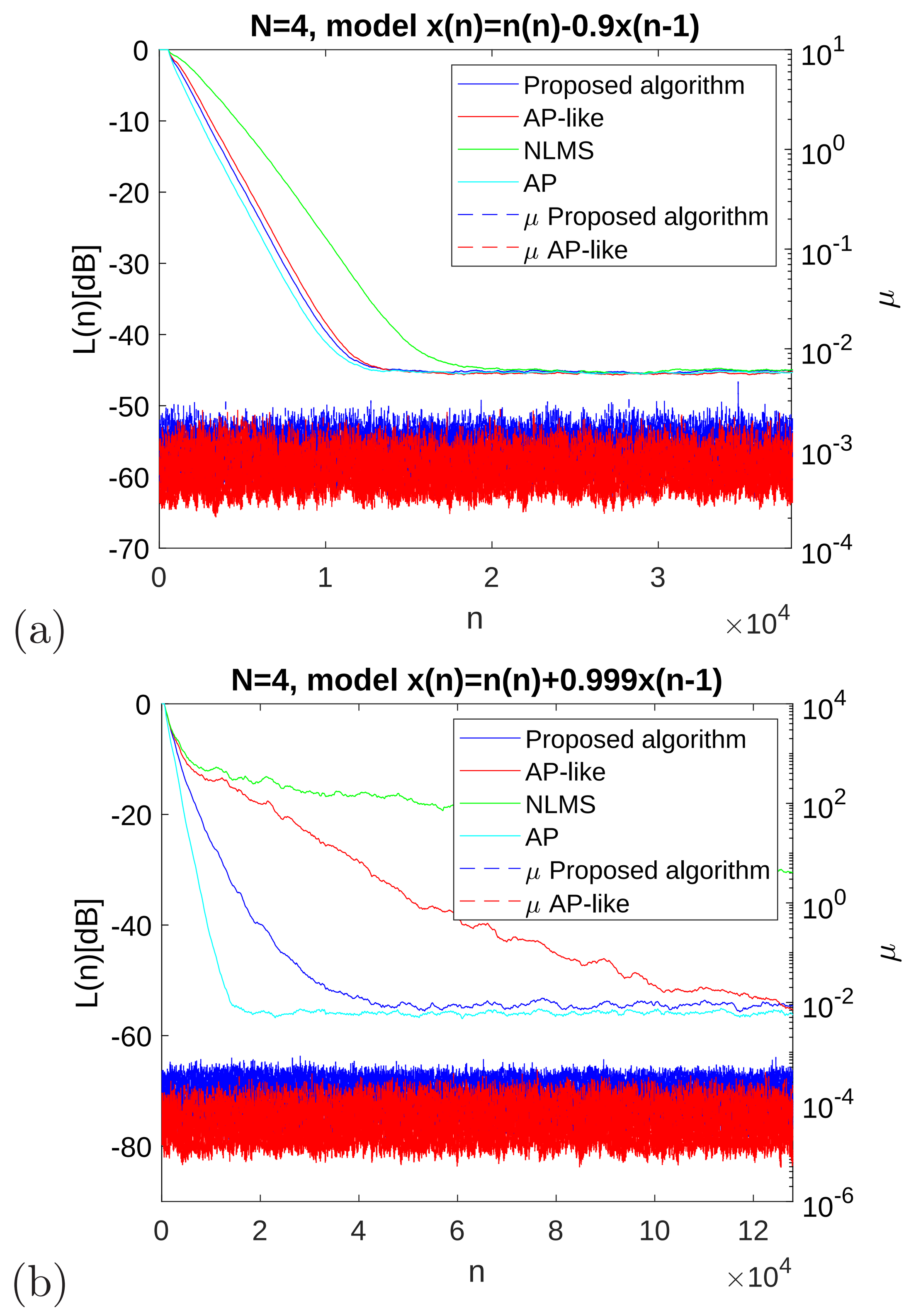

Figure 2: Comparative learning curves and step-size values for: AP, APL-I and the proposed algorithm for $N = 4$ when the input signal is: (a) slightly colored ($\gamma = 0.9$) and (b) highly colored ($\gamma = 0.999$).

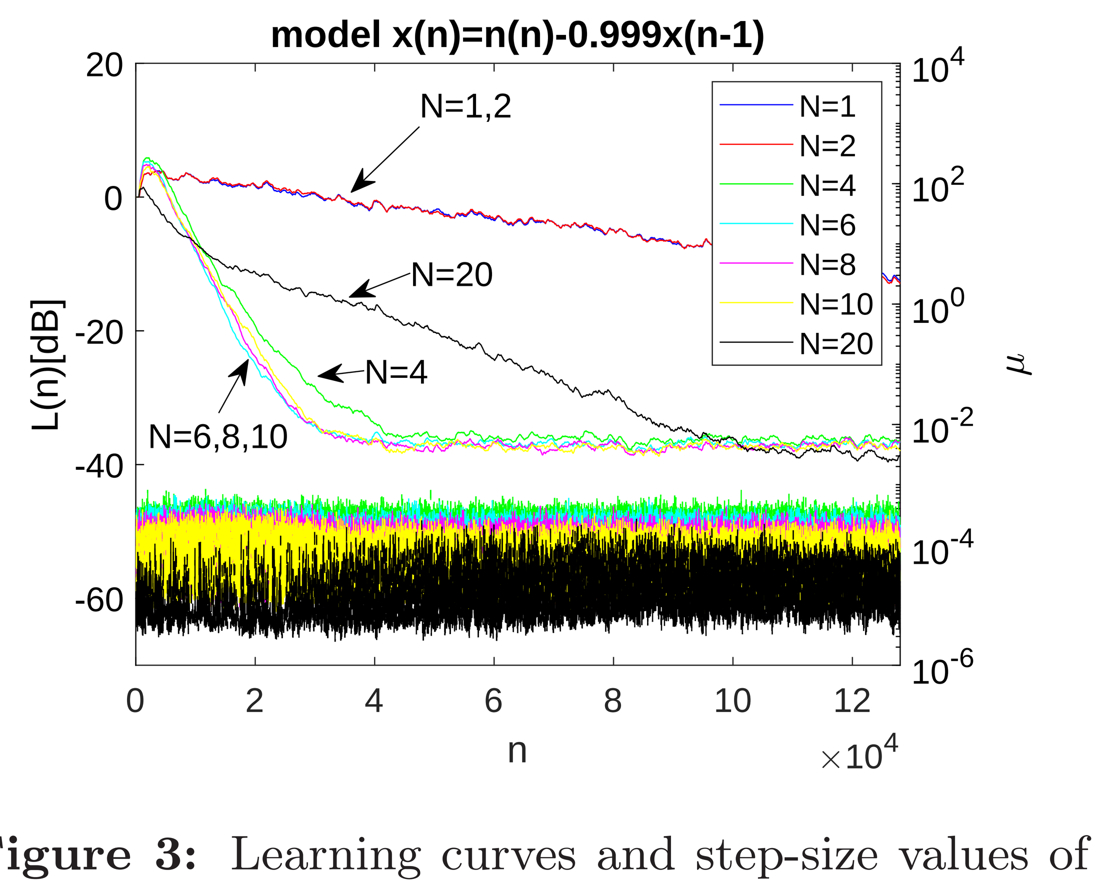

Figure 3: Learning curves and step-size values of the proposed algorithm and different projection orders for highly colored reference signal.