AN ADAPTIVE COMBINATION OF ADAPTIVE FILTERS FOR PLANT IDENTIFICATION

M. Martínez-Ramón, J. Arenas-García, A. Navia-Vázquez and A. R. Figueiras-Vidal

Department of Signal Theory and Communications
Universidad Carlos III de Madrid
Avda. de la Universidad, 30 / 28911 Leganés (Madrid), Spain
arfv@tsc.uc3m.es

Abstract: In this paper, we propose to adaptively combine two LMS adaptive transversal filters for plant identification. One of the filters has a high and the other a low adaption steps, in order to combine good tracking capabilities under (fast) change conditions with a reduced convergence error along stationary periods. A brief discussion of the characteristics of the combination is included, emphasizing that it allows the possibility of dealing with "intermediate" rate of change situations, in opposition to (implicit or explicit) switching mechanisms. A selected illustrative simulation example shows the effectiveness of this approach. Some complementary lines of research are indicated in the Conclusions, from the points of view of improving the algorithm and of extending the fields of application.

1. INTRODUCTION

Widrow and Hoff's Least Mean Square (LMS) [1] algorithm is a straightforward and robust alternative for adaptive transversal filtering, including system identification, as recognized in many excellent textbooks [2,3,4]. A good compromise between the stationary error level and its tracking capabilities is probably the main advantage of using LMS. The adaptation step serves to select the exact balance between speed and convergence.

So, it becomes an evidence that the standard LMS has just a parameter –the adaptation step– to establish the above trade-off. It is also obvious that a little more complexity can be accepted if the degrees of freedom are increased, in order to improve the speed vs. convergence balance.

A series of modifications have appeared in the literature along this direction: some examples in system identification are the work of Harris et al. [5], in which a change in the values of the adaptation step according to the recent error signs allows to modify the trade off according to the situation[1]; some time-varying combinations of LMS and Least Mean Fourth (LMF, [6]) algorithms by means of several empirical rules [7,8], that try to take advantage of the different characteristics of both types of algorithms; and even a combination of LMS and the Least Mean Absolute algorithm to combat the presence of impulsive noise in plant identification [9].

In all the above cases but the first, the designs combine different costs to obtain a change in the speed vs. convergence compromise. In [5], just the LMS algorithm adaptation step is increased/decreased according to the convenience of getting higher speed/lower error, respectively.

We consider that the underlying idea of using an adaptive step is appropriate. However, it seems clear that to design the corresponding adaptation rule is a delicate question, indeed. In particular, to use error signs opens the door to suffer the negative effects of noise and non-monotonic convergence: it is a well known fact that the error has very "noisy" registers in these schemes.

Here, we follow the above basic idea, but under a completely different approach.

2. THE PROPOSED SCHEME

We propose to use for plant identification an adaptive convex combination of two LMS adaptive transversal filters. Obviously, this is equivalent to just one transversal filter at each time instant. One of the adaptive filters, $\mathbf{w}_1[k]$, is "fast": its adaptation constant, μ_1, is (relatively) high. The second plant, $\mathbf{w}_2[k]$, is "slow": it works with a (relatively) low adaptation step, μ_2. But this second plant would provide, in steady-state situations, a lower plant identification error.

So, if we use

$$\mathbf{w}_{eq}[k] = \lambda[k]\,\mathbf{w}_1[k] + \big(1 - \lambda[k]\big)\mathbf{w}_2[k] \qquad (1)$$

and we force that $\lambda[k]$ be high (near 1, if needed) for fast tracking situations, and low (near 0) for stationary intervals, we are combining the higher tracking capabilities of $\mathbf{w}_1[k]$ with the better steady-state performance of $\mathbf{w}_2[k]$.

To do so, it is enough to adapt λ according to an LMS algorithm with an adaption parameter μ_λ much higher than μ_1 (and, consequently, μ_2). There is not any

[1] In fact, [5] proposes to do so for each weight component, thus alleviating the well known convergence difficulties originated by the eigenvalue spread: but, in our context, the above interpretation is adequate.

significant divergence problem: $\lambda[k]$ is kept into [0,1], to get just a convex combination (in other case, the scheme can go out of control).

The way of adapting λ comes from

$$\nabla_\lambda e^2 = \nabla_\lambda \left(y - \mathbf{w}_{eq}^T \mathbf{x}\right)^2 = \nabla_\lambda \left(y - \left\{\lambda \mathbf{w}_1^T + (1-\lambda)\mathbf{w}_2^T\right\}\mathbf{x}\right)^2$$
$$= -2\left(y - \mathbf{w}_{eq}^T\mathbf{x}\right)\left(\mathbf{w}_1^T - \mathbf{w}_2^T\right)\mathbf{x} \quad (2)$$

So, and according to Fig. 1, we apply the following algorithms: standard LMS for the filters

$$\mathbf{w}_i[k+1] = \mathbf{w}_i[k] + \mu_i \left(y[k] - \mathbf{w}_i^T[k]\,\mathbf{x}[k]\right)\mathbf{x}[k] \quad i = 1,2 \quad (3a)$$

and, according to (2), the same for the convex combination parameter

$$\lambda[k+1] = \lambda[k] + \mu_\lambda \left(y[k] - \mathbf{w}_{eq}^T[k]\,\mathbf{x}[k]\right)$$
$$\left(\mathbf{w}_2^T[k] - \mathbf{w}_1^T[k]\right)\mathbf{x}[k]\left\{\lambda[k]\left(1 - \lambda[k]\right) + 10^{-2}\right\} \quad (3b)$$
$$\text{keeping} \quad 0 \le \lambda[k] \le 1$$

and (1) to estimate $\mathbf{w}[k]$, the unknown transversal plant; $\mathbf{x}[k]$ and $y[k]$ being the input vector and the noisy output of this plant, respectively ($y[k] = \mathbf{w}^T\,\mathbf{x}[k] + n[k]$, n being the additive noise).

Note that we add a factor $\lambda[k](1-\lambda[k])+10^{-2}$ to the λ gradient: just to reduce the speed of convergence for λ near to 0 or 1, without stopping the algorithm.

A qualitative discussion of the performance of this scheme is very simple: it tends to have the tracking capabilities corresponding to μ_1 ($\lambda[k] \to 1$) for very abrupt changes, and the residual error corresponding to μ_2 ($\lambda[k] \to 0$) for long stationary periods. In intermediate situations, $\lambda[k]$ will take intermediate values, and the convergence will be a "mixture" of that of $\mathbf{w}_1[k]$ and $\mathbf{w}_2[k]$. So, we get a reasonable combination of both LMS algorithms, assuming that the combination is not very "noisy". It is also clear that the usual convergence conditions must be accomplished by μ_1.

Let us remark that this approach has the additional advantage of avoiding "switching" procedures: to apply one of the filters at each moment is not an appropriate solution when the rate of change is not extreme (abrupt, or at least very fast, or zero, or at least very low). In some form, most of the previously proposed schemes suffer this drawback.

3. AN ILLUSTRATIVE EXPERIMENT

We have carried out a large number of experiments, that support both the efectiveness of the proposed scheme and its interpretation. Among them, we select a simulation example that shows all the main aspects that need to be considered.

In our example, x and n are independent, white, Gaussian, zero-mean noises with variances 1 and 10^{-2}, respectively, while

$$\mathbf{w}[k] = \mathbf{w}_C + \mathbf{w}_V[k] \quad (4)$$

is a 3-tap plant, where

$$\mathbf{w}_C = \left[-0.4706, -0.7737, -0.0291\right]^T$$

for $1 \le k \le 21000$, and

$$\mathbf{w}_C = \left[0.4706, 0.7737, 0.0291\right]^T$$

along $21001 \le k \le 28000$ (the coefficient values have been selected at random from a [0,1) uniform distribution), and $\mathbf{w}_v[k]$ is the result of decimating 436:1 plus optimally interpolating a noise sequence as the above ones, but with variances

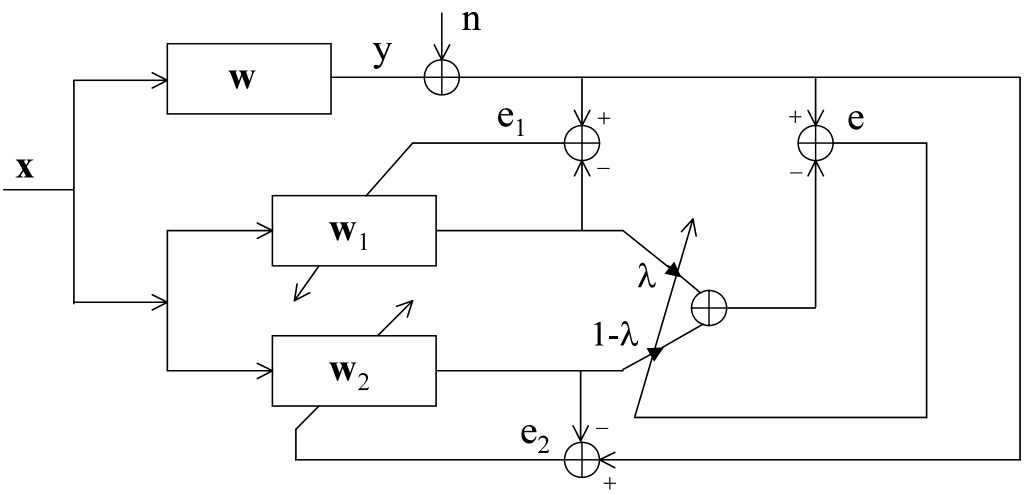

Fig 1: The proposed adaptive (convex) combination scheme. Two LMS adaptive filters, one of them "fast" and the other "slow" (high and low adaption step, respectively), adaptively combine their outputs by means of a very fastly adapted parameter λ.

$$
\begin{array}{lrcl}
10^{-4}, & 0 & \le k \le & 7000 \\
10^{-2}, & 7001 & \le k \le & 14000 \\
0 \text{ (no noise)}, & 14001 & \le k \le & 21000 \\
10^{-2}, & 21001 & \le k \le & 28000
\end{array} \quad (5)
$$

This plant has been selected to show the performance of the combined scheme for fast and slow changes as well as during stationary periods.

The parameters of the algorithm are: $\mu_1 = 0.05$; $\mu_2 = 0.002$; $\mu_\lambda = 20$.

Fig. 2 shows the average of 1000 experiments: 2(a) shows (in dB) the average of the squared distance between the unknown plant and the adaptative plant

$$\mathrm{MSD}[k] = \left\| \mathbf{w}[k] - \mathbf{w}_1[k] \right\|_2^2 \quad (6)$$

for the application of $\mathbf{w}_1[k]$, and Fig. 2(b) and 2(c), the same magnitude when applying $\mathbf{w}_2[k]$ and the proposed $\mathbf{w}_{eq}[k]$, respectively.

As expected, the filter with the high step, $\mu_1 = 0.05$, shows its better tracking capabilities at the beginning, during the fast change interval $7001 \le k \le 14000$, and when the abrupt plant change occurs ($k = 21001$); however, it is unable of reaching an error below –30 dB during the quiet interval $1 \le k \le 7000$. On the contrary, the filter with adaption step $\mu_2 = 0.002$ is slower when the abrupt changes occur, but it provides also lower error: not only for the quiet interval $1 \le k \le 7000$, but also for the final segment $21001 \le k \le 28000$; although during $7001 \le k \le 14000$ it is unable of following the plant changes.

It can be seen that the combined scheme has the tracking capabilities of the fast LMS (see the initial convergence and that after the jump at $k = 21001$, as well as its performance during $7001 \le k \le 14000$, while $\mathbf{w}_2[k]$ is unable to track the plant), together with the low error of the "slow" filter for stationary periods ($14001 \le k \le 21000$).

To support our point of view about the mechanics of our scheme, we show the (averaged) evolution of $\lambda[k]$ for our example in Fig. 3.

4. CONCLUSIONS

A direct adaptive convex combination of one fast and one slow adaptive filters allows to get both the tracking capabilities of the first and the high degree of convergence of the later for plant identification if the adaptation of the combination is fast enough. The work of the combined scheme is easily understandable, and so is its design. Application examples support the usefulness of the approach and its interpretation.

There are, of course, many possibilities to extend the proposed ideas: the first, to explore other reference errors for $\mathbf{w}_1$ and $\mathbf{w}_2$. Other extensions go from using a

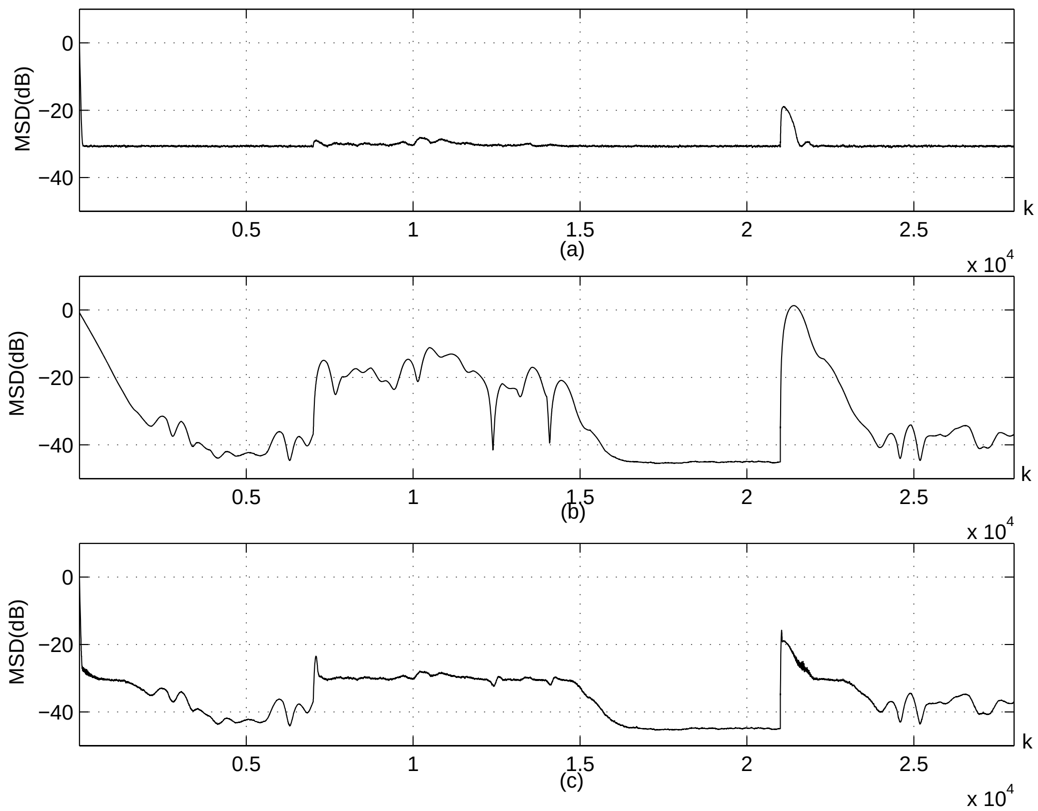

Fig 2: Plant Identification error for the example when using: (a) A "fast" LMS ($\mu_1 = 5\cdot10^{-2}$); (b) A "slow LMS ($\mu_2 = 2\cdot10^{-3}$); (c) The adaptively combined algorithm (μ_1 and μ_2 as above, $\mu_\lambda = 20$).

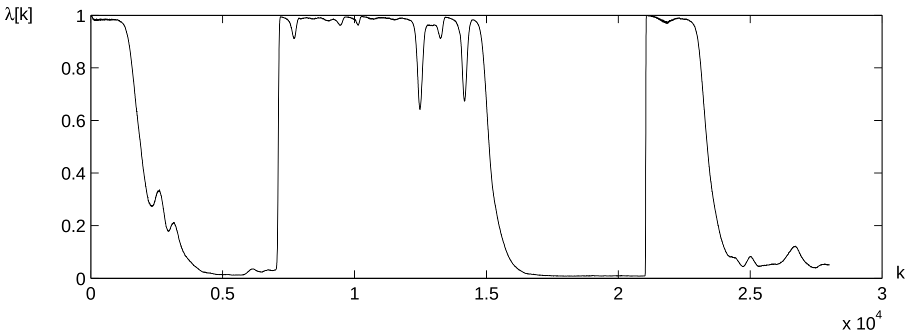

Fig 3: (Averaged) evolution of λ for the example.

variable-step μ_λ depending on the own $\lambda[k]$ values (note that these values are 'smoother' than the errors when we approach the 0 or 1 limits), improving again the capacities of the basic scheme, to include a diagonal matrix $\Lambda[k]$ with differently adapted elements in order to combat eigenvalue dispersion, in a way similar to that in [5] but keeping the advantages of our model.

Needless to say, it is also possible to apply these concepts to other adaptive filtering problems, going from digital equalization to adaptive arrays, just by taking the adequate providences to build "reasonable" designs.

All these extensions are currently being explored by the authors.

REFERENCES

[1] B. Widrow, M. E. Hoff, "Adaptive Switching Circuits", *Wescon Conv. Record*, Pt. 4, pp. 96-140, 1960.

[2] S. Haykin, *Adaptive Filter Theory (4th Ed.)*, Upple Saddle River, NJ: Prentice-Hall, 2002.

[3] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, New York, NY: Wiley, 1996.

[4] N. Kaloupsidis, S. Theodoridis, *Adaptive System Identification and Signal Processing Algorithms*, Hemel Hempstead, Heartfordshire, UK: Prentice-Hall, 1993.

[5] R. W. Harris, D. M. Chabries, F. A. Bishop, "Variable Step (VS) Adaptive Filter Algorithm", *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-34, pp. 305-316, 1986.

[6] E. Walach, B. Widrow, "The Least Mean Fourth (LMF) Algorithm and Its Family", *IEEE Trans. Information Theory*, vol. IT-30, nº 2, Pt. 1, pp. 275-283, 1984.

[7] D. I. Pazaitis, A. G. Constantinides, "LMS + F Algorithm", *Electron. Lett.*, vol. 31, nº 17, pp. 1423-1424, 1995.

[8] C. Rusu, M. Helsingius, P. Kuosmanen, "Joined LMS and LMF Threshold Technique for Data Echo Cancellation", in *Proc. COST#254 Intl. Workshop on Intelligent Comms. And Multimedia Terminals* (J. Tasic, ed.), pp. 127-130, Ljubliana, Slovenia, November 1998.

[9] J. Chambers, A. Avionitis, "A Robust Mixed-Norm Adaptive Filter Algorithm", *IEEE Signal Proc. Letters*, vol. 4, pp. 46-48, 1997.

Acknowledgments

This work has been partly supported by grants CICYT TIC 1999-0216 and CAM 07T/0046/2000.