International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:8, No:7, 2014

Combination Scheme of Affine Projection Algorithm Filters with Complementary Order

Young-Seok Choi

***Abstract*—**This paper proposes a complementary combination scheme of affine projection algorithm (APA) filters with different order of input regressors. A convex combination provides an interesting way to keep the advantage of APA having different order of input regressors. Consequently, a novel APA which has the rapid convergence and the reduced steady-state error is derived. Experimental results show the good properties of the proposed algorithm.

***Keywords*—**Adaptive filter, affine projection algorithm, convex combination, input order.

I. Introduction

THE normalized least mean square (NLMS) is one of widely used adaptive algorithms due to its simplicity and ease of implementation. However, its convergence rate is significantly reduced for correlated input signals [1]-[3]. To overcome this problem, the affine projection algorithm (APA) was proposed by Ozeki and Umeda [2]. While the LMS-type filters update the weights based only on the current input vectors, the APA updates the weights on the basis of the last K input vectors. Also it is known [2] that for the APA the number of input regressors plays a critical role in the convergence performance of the APA. In the APA, the number of input regressors governs both the rate of convergence and the steady-state error. To meet the conflicting requirements of fast convergence and low steady-state error, the number of input regressors needs to be optimized. Fig. 1 shows the mean square deviation (MSD) of the APA when the number of input regressors is different.

Recently, to improve the performance of the APA, several algorithms which focus on the number of input regressors have been presented [4],[5]. The selective regressor APA(SR-APA) [4] focuses on selecting a subset of a fixed number of members of input regressors at each time. This algorithm shows us that each input regressor vector has a different influence to the convergence performance. In addition, set-membership APA with variable data-reuse factor which based on the set-membership APA has presented that varying the number of the input regressors can improve the convergence performance [5].

In the adaptive filtering literature, an approach consisting on an adaptive convex combination of two adaptive filters: one of them is fast, the other is slow [6], [7]. These filters are combined in such manner that the advantages of both component filters are kept: the rapid convergence from the fast filter, and the reduced steady-state error from the slow filter.

Y.-S. Choi is with the Department of Electronic Engineering, Gangneung-Wonju National University, 7 Jukheon-gil, Gangneung, 210-702 Republic of Korea (phone: +82-33-640-2429; fax: +82-33-646-0740; e-mail:yschoi@gwnu.ac.kr).

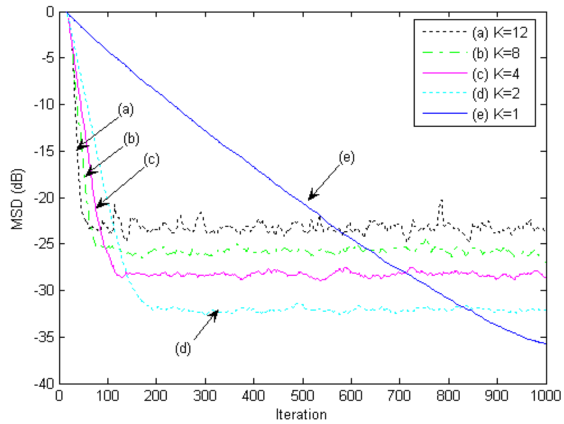


Fig. 1 Plots of MSD for the APA with different input order

With this in mind, we propose an adaptive convex combination of the APA filters with different order of input regressors. The proposed combination of the APA (C-APA) dynamically adjusts the mixing parameter to combine two APA filters using a stochastic gradient method. The resulting scheme is able to extract the best properties of each component filter, namely fast convergence after abrupt changes in the channel and low residual error in steady-state.

This paper is organized as follows. In Section II, we introduce the conventional APA and develop the C-APA. Section III illustrates the experimental results and Section IV present conclusion.

II. Convex Combination of APA filters

Consider data $d(i)$ that arise from the system identification model

$$d(i) = \mathbf{u}_i \mathbf{w}^o + v(i), \tag{1}$$

where $\mathbf{w}^o$ is a column vector for the impulse response of an unknown system that we wish to estimate, $v(i)$ accounts for measurement noise and $\mathbf{u}_i$ denotes the $1 \times M$ row input vector,

$$\mathbf{u}_i = [u(i)\ u(i-1) \cdots u(i-M+1)], \tag{2}$$

and $\mathbf{u}_i$ and $v(i)$ are uncorrelated.

A. Conventional APA

Let $\mathbf{w}_i$ be an estimate for $\mathbf{w}^o$ at iteration i. The conventional APA computes $\mathbf{w}_i$ via

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mu \mathbf{U}_i^* (\mathbf{U}_i \mathbf{U}_i^*)^{-1} \mathbf{e}_i, \tag{3}$$

where

$$\mathbf{U}_i = \begin{bmatrix} \mathbf{u}_i \\ \mathbf{u}_{i-1} \\ \vdots \\ \mathbf{u}_{i-K+1} \end{bmatrix}, \quad \mathbf{d}_i = \begin{bmatrix} d(i) \\ d(i-1) \\ \vdots \\ d(i-K+1) \end{bmatrix}, \quad \mathbf{e}_i = \mathbf{d}_i - \mathbf{U}_i \mathbf{w}_{i-1},$$

μ is the step-size, and * denotes the Hermitian transpose. It is well known that for the convergence performance of the APA depends on the number of input regressors to be used for tap weight adaptation, i.e., there exists a tradeoff between the filter convergence rate and the steady-state error according to the number of input regressors [3]. If a larger number of the input regressors are used, then a faster convergence is attained while indicating a higher steady-state error. On the other hand, the smaller number of input regressors the APA takes, the more accurate the estimation is obtained with the poorer convergence rate. Along this line of thought, we may expect performance improvement through solving this tradeoff according to the input regressors.

III. Convex Combination of APA Filters

Our objective is to improve the convergence performance by using an adaptive convex combination of two APA filters, the first being a fast filter (i.e., with a large number of input regressors) and the second a slow filter (with a small number of input regressors). The output signals and the output errors of both filters are combined to obtain advantage of each filter: the rapid convergence from the fast filter and the reduced steady-state error form the slow filter. We assume the first filter as a fast filter.

We start with the following adaptive convex combination of APA scheme, which obtains the output of the overall filter as

$$y(i) = \lambda(i) y_1(i) + [1-\lambda(i)] y_2(i), \tag{4}$$

where $y_1(i)$ and $y_2(i)$ are the outputs of two APA filters at time i, i.e., $y_k(i) = \mathbf{u}_i \mathbf{w}_{k,i-1}, k = 1,2$, $\mathbf{w}_{k,i-1}$ is the adaptive filter weight vector of k th filters, and $\lambda(i) \in [0\ 1]$ is a mixing scalar parameter. By choosing appropriate mixing parameter $\lambda(i)$ at each time, the combination scheme (4) would result in the respective merits of both APA filters. Therefore, if both $\mathbf{w}_{1,i-1}$ and $\mathbf{w}_{2,i-1}$ have same length, then overall filter is obtained by

$$\mathbf{w}_{i-1} = \lambda(i) \mathbf{w}_{1,i-1} + [1-\lambda(i)] \mathbf{w}_{2,i-1}. \tag{5}$$

TABLE I
Proposed C-APA

Initialization:
$\mathbf{w}_{-1} = \mathbf{0}$
Vector definition:
$\mathbf{u}_i = [u(i)\ u(i-1)\ \cdots u(i-M+1)]$
For each new input $u(i);\ i \geq 0$:
$\mathbf{w}_{k,i} = \mathbf{w}_{k,i-1} + \mu \mathbf{U}_{k,i}^* (\mathbf{U}_{k,i} \mathbf{U}_{k,i}^*)^{-1} \mathbf{e}_{k,i}$
$\mathbf{e}_{k,i} = \mathbf{d}_i - \mathbf{U}_{k,i} \mathbf{w}_{k,i-1},\ k = 1,2$
$a(i) = a(i-1) + \mu_a e(i)[y_1(i-1) - y_2(i-1)] \cdot \lambda(i-1)[1-\lambda(i-1)]$
$\lambda(i) = \mathrm{sgm}[a(i)] = (1+e^{-a(i)})^{-1}$
$\mathbf{w}_i = \lambda(i) \mathbf{w}_{1,i} + [1-\lambda(i)] \mathbf{w}_{2,i}$

In the APA literature, both APA filters are adapted independently using their own output as

$$\mathbf{w}_{k,i} = \mathbf{w}_{k,i-1} + \mu \mathbf{U}_{k,i}^* (\mathbf{U}_{k,i} \mathbf{U}_{k,i}^*)^{-1} \mathbf{e}_{k,i}, \tag{6}$$

where $\mathbf{e}_{k,i} = \mathbf{d}_i - \mathbf{U}_{k,i} \mathbf{w}_{k,i-1},\ k = 1,2$.

For the determination of the mixing parameter $\lambda(i)$, we incorporated a stochastic gradient method so that $J(i) = e^2(i)$ is minimized where $e(i) = d(i) - \mathbf{u}_i \mathbf{w}_{i-1}$. However, rather than adapting $\lambda(i)$ directly, we will update a new parameter $a(i)$ which defines $\lambda(i)$ through the sigmoidal function as

$$\lambda(i) = \mathrm{sgm}[a(i)] = (1+e^{-a(i)})^{-1}. \tag{7}$$

Accordingly, the update equation for $a(i)$ is given by

$$\begin{aligned} a(i) &= a(i-1) - \frac{\mu_a}{2} \cdot \frac{\partial J(i)}{\partial a(i-1)} \\ &= a(i-1) - \frac{\mu_a}{2} \cdot \frac{\partial e^2(i)}{\partial \lambda(i-1)} \cdot \frac{\partial \lambda(i-1)}{\lambda a(i-1)} \\ &= a(i-1) + \mu_a e(i)[y_1(i-1) - y_2(i-1)] \\ &\quad \cdot \lambda(i-1)[1-\lambda(i-1)], \end{aligned} \tag{8}$$

where μ_a is the step-size for the adaptation of $a(i)$ and should be chosen to a very high value to guarantee the fast adaptation of the combination over the fast APA filter.

By introducing the sigmoid function, we can easily guarantee $\lambda(i) \in [0\ 1]$. In the practical application, $a(i)$ is limited to the interval [–4 4] to forbid the algorithm to step when $\lambda(i)$ is too close 0 or 1 [6].

The performance of the combination algorithm can be improved by two modification scheme [6]. One of them is to speed up the convergence of the slow filter by transferring the advantage of fast one and is given by

$$\mathbf{w}_2(i) = \alpha \mathbf{w}_2(i) + (1-\alpha)\mathbf{w}_1(i). \quad (9)$$

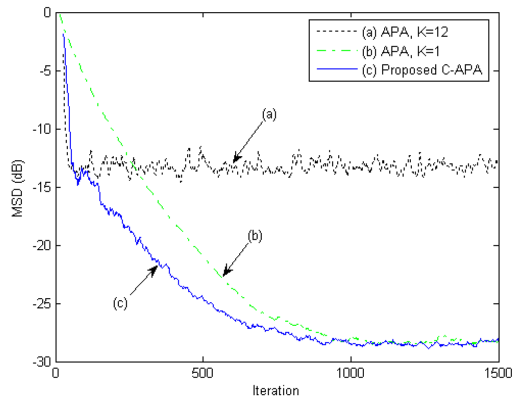
where α is a parameter close to 1.

Another modification is to improve the convergence of the parameter $a(i)$. In (8), when $y_1(i-1) - y_2(i-1)$ is very small, the adaptation of is deteriorated. we can alleviate this problem by inserting the momentum term as the follows

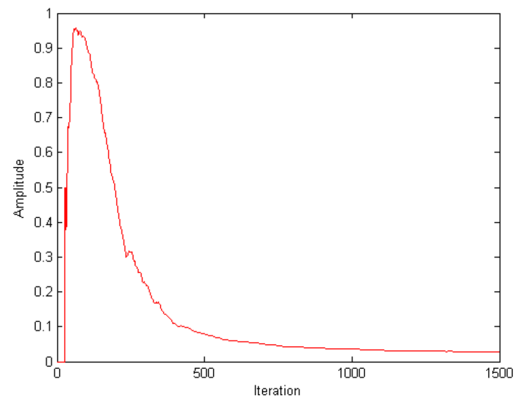
$$a(i) = a(i-1) + \mu_a e(i)[y_1(i-1) - y_2(i-1)] \cdot \lambda(i-1)[1-\lambda(i-1)] + \rho[a(i) - a(i-1)]. \quad (10)$$

where ρ is a positive constant. We incorporate these two modification schemes in the experiment.

The proposed C-APA with different number of input repressors has the following feature. For initial period or fast change, the fast APA filter is superior to the slow one making $\lambda(i)$ be toward 1. On the contrary, in stationary period, the slow APA filter shows the better performance, which results in $\lambda(i)$ approximated as 0.



(a)



(b)

Fig. 2 (a) Plots of MSD for the APA and the proposed C-APA (b) Plot of the mixing parameter $\lambda(i)$

IV. Experimental Results

We illustrate the performance of the proposed algorithm by carrying out computer simulations in a channel identification scenario. The unknown channel $H(z)$ has 16 taps and is randomly generated. The adaptive filter and the unknown channel are assumed to have the same number of taps. The input signal is obtained by filtering a white, zero-mean, Gaussian random sequence through a first-ordersystem, which is given by

$$G(z) = 1/(1-0.9z^{-1}).$$

The signal-to-noise ratio (SNR) is calculated by

$$\text{SNR} = 10\log_{10}\left(\frac{E[y^2(i)]}{E[v^2(i)]}\right),$$

where $y(i) = \mathbf{u}_i \mathbf{w}^o$. The measurement noise $v(i)$ is added to $y(i)$ such that SNR=30dB. The meansquare deviation (MSD), $E\|\mathbf{w}^o - \mathbf{w}_i\|^2$, is taken and averaged over 100 independent trials. The step-size is set as $\mu = 0.5$. To adapt the combination, we use $\mu_a = 100$. And $\alpha = 0.9$ and $\rho = 0.5$ are used. Finally, the mixing coefficient $a(i)$ has been initialized to zero.

Fig. 2 (a) shows the MSD curves of the conventional APA and the proposed C-APA. Dashed lines indicate the results of the APA with $K = 12$ and $K = 1$ (identical to NLMS). As can be seen, the proposed C-APA inherits the best properties of each component filter, presenting the fast convergence rate with the low steady-state error comparable to $K = 12$ and $K = 1$, respectively.

In Fig. 2 (b), the evolution of the mixing parameter $\lambda(i)$ is shown. We can see that the parameter $\lambda(i)$ increase toward unity initially, to provide the fast convergence rate for the overall filter and then converges to a small value to obtain a small steady-state error from the slow filter.

V. Conclusion

We have proposed a convex combination of the APA filters which has different order of input regressors. By adjusting the mixing parameter dynamically, two APA filters can be combined into overall filter. The proposed C-APA is able to extract the best properties of each component filter, namely the fast convergence rate and the low steady-state error.

Reference

[1] S. Haykin, *Adaptive Filter Theory*, 4th edition, Upper Saddle River, NJ: Prentice Hall, 2002.
[2] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electro. Commun. Jpn.*, vol. 67-A, no. 5, pp. 19-27, 1984.
[3] H.-C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Processing*, vol. 52, pp. 90-102, Jan. 2004.
[4] K.-Y. Hwang, W.-J. Song, "An affine projection adaptive filtering algorithm with selective regressors," *IEEE Trans. Circuits and Syst. II.*,

vol. 54, pp. 43-46, Jan.. 2007.
[5] S. Werner, P. S. R. Diniz, and J. E. W. Moreira, "Set-membership affine projection algorithm with variable data-reuse factor," in *Proc. 2006 IEEE Int. Wymp. Circuits and Systems*, pp. 261-264, May 2006.
[6] J. Arenas-García, V. Gómez-Verdejo, and A. R. Figueiras-Vidal, "New algorithms for improved adaptive convex combination of LMS transversal filters," *IEEE Trans. Instrum. Meas.*, vol. 54, pp. 2239-2249, Dec. 2005.
[7] J. Arenas-García, A. R. Figueiras-Vidal, and A. H. Sayed, "Steady-state performance of convex combinations of adaptive filters," in *Proc. ICASSP*, Philadelphia, PA, May 2005.