

Absolute Mean Error Based Adaptive Step Size Affine Projection Algorithm for Adaptive Filtering Systems

Thamer M. Jamel

Abstract—This paper presents a new method for adaptive step size for Affine Projection Algorithm (APA). The proposed algorithm is based on an absolute mean of estimation current and prior error vector. It is called Absolute Mean Error Adaptive Step Size Affine Projection Algorithm (AMASSAPA). The main goal of this algorithm is performance enhancement of adaptive colored filtering system in terms of fast convergence time. The proposed adaptive step size method begins the learning process with high learning rate value (μ_{MAX}) and then, it decays in an exponential profile to its minimum value (μ_{MIN}). The proposed algorithm is tested with a colored input signal and the result shows that it has fast convergence time than traditional APA.

Index Terms—Adaptive filter, adaptive step size, affine projection algorithm.

I. INTRODUCTION

The most common adaptive algorithms is Least Mean Square (LMS) algorithm, but this algorithm suffers from slow convergence speed if it has driven by colored input signals as is with speech. One method presented to overcome this problem is the Ozeki/ Umeda Affine Projection Algorithm (APA) [1]-[4]. APA is a useful family of adaptive filters whose main purpose is to accelerate the convergence of LMS-type filters, especially for correlated data at a computational cost that is comparable to that of LMS. While LMS updates the weights based on the current input vector and APA updates the weights based on previous input vectors [5]. The algorithm applies to update directions that are orthogonal to the last p input vectors and thus allows decorrelation of an input process, speeding up the convergence [2].

An APA with a constant step-size parameter has to compromise between the performance's criteria of fast convergence rate, and low misadjustment. Therefore, a variable step-size APA represents a more reliable solution [6].

In order to meet this conflicting requirement, a number of variable step-size APAs were developed [1], [3], [6], [7], [8], [9], [10] and [11]. Tao Dai, Andy Adler, and Behnam Shahrava [1] present a forgetting factor scheme for variable step-size affine projection algorithms (APA). The proposed scheme uses a forgetting processed input matrix as the projection matrix of pseudo-inverse to estimate system deviation. Mohammad Shams, Vahid Mehrdad, Azizollah Gholipour and Majid Noroozi [3] established family of VSS

selective partial updates (SPU) affine projection algorithms which reduce the computational complexity. Constantin Paleologu, Jacob Benesty, and Silviu Ciochin'a [6], derived a VSS-APA in the context of acoustic echo cancellation (AEC). The proposed VSS-APA aims to recover the near-end signal within the error signal of the adaptive filter. F. Albu, C. Paleologu [7] proposed, a new pseudo affine projection (AP) algorithm based on Gauss-Seidel (GS) iterations for acoustic echo cancellation (AEC). Aditya and Daljit [8] proposed a novel variable step-size APA (VSS-APA) algorithm for multiuser detection. Ligang Liu, Masahiro Fukumoto, Sachio Saiki, and Shiyong Zhang [9], proposed a variable step-size proportionate affine projection algorithm (PAPA) by canceling the a posteriori estimation error.

Murat Cubak [10] proposed Gradient Adaptive Step Size method for APA. The GASAPA provides both fast convergence associated with zero forcing and minimum misadjustment associated with the optimum step size. Kirill Sakhnov [11] proposed proportionate normalized least mean squares (PNLMS). The idea behind PNLMS was to update each coefficient of the adaptive filter independently of the others by adjusting the adaptation step size in proportion to the magnitude of the estimated filter coefficient.

Adjusting step size instead of fixed step size has a powerful effect on the performance of the system and the structure of the adaptive filter will not be changed, also this technique requires fewer overheads in computations, which are an important factor for hardware implementation. In this paper, a new adaptive step size method for APA is proposed, which is called AMASSAPA. The main goal of this proposed algorithm is achieving fast convergence time when colored input signal was used in the adaptive filtering system.

This paper is organized as follows: Section 2 presents the basic concept of APA, while section 3 gives the proposed algorithm AMASSAPA. Section 4 demonstrates the performance of the proposed algorithm by carrying out experiments in system identification and finally section 5 gives important conclusions about this research.

II. AFFINE PROJECTION ALGORITHM (APA)

In this section the basic concept of the APA will be introduced for the adaptive filter system shown in Fig. 1. The input vector $x(k)$ can be defined as:

$$x(k) = [x(k), x(k-1), \dots, x(k-L+1)]^T \quad (1)$$

where k is the discrete time index, T denotes to transpose, and adaptive Finite Impulse Response (FIR) filter coefficient vector defined by [12]

Manuscript received July 27, 2012; revised August 28, 2012.
Thamer M. Jamel is with the Department of Electrical Engineering, Technology University, Baghdad, Iraq (e-mail: dr.thamerjamel@uotechnology.edu.iq)

$$\mathbf{w}(k) = [w_1(k), w_2(k), \dots, \dots, w_L(k)]^T \quad (2)$$

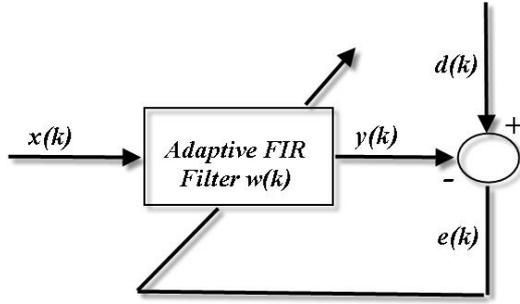


Fig. 1. Block diagram of adaptive filtering system

where $w_1(k)$, $w_2(k)$, ..., $w_L(k)$ are filter coefficients at time k , L is the number of taps (or coefficients). The output of the adaptive filter (k) is calculated by

$$y(k) = \mathbf{x}^T(k) \mathbf{w}(k) \quad (3)$$

The objective of an adaptive filter is to generate the output signal $y(k)$ as close as possible to the desired signal $d(k)$. To do this, the adaptive filter adjusts its coefficient $\mathbf{w}(k)$ at each sampling time k by [12]

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \Delta \mathbf{w}(k) \quad (4)$$

where $\Delta \mathbf{w}(k)$ is an adjustment vector and μ is called the step size, which controls the amount of adjustments. The adjustment vector $\Delta \mathbf{w}(k)$ is determined in the projection algorithm to satisfy the following equations [13]:

$$d(k) = \mathbf{x}^T(k) (\mathbf{w}(k) + \Delta \mathbf{w}(k))$$

$$d(k-1) = \mathbf{x}^T(k-1) (\mathbf{w}(k) + \Delta \mathbf{w}(k))$$

.....

$$d(k-p+1) = \mathbf{x}^T(k-p+1) (\mathbf{w}(k) + \Delta \mathbf{w}(k)) \quad (5)$$

where p ($p < L$) is called the projection order. These equations mean that adjusted vector $\mathbf{w}(k+1)$ with $\mu = 1$ (that is $\mathbf{w}(k) + \Delta \mathbf{w}(k)$) generates $d(k)$, ..., $d(k-p+1)$ for the past p input vectors $x(k)$, ..., $x(k-p+1)$. Equation (5) can be rewritten in matrix form as [13]:

$$\mathbf{d}_p(k) = \mathbf{X}_p^T(k) (\mathbf{w}(k) + \Delta \mathbf{w}(k)) \quad (6)$$

or equivalently, as

$$\mathbf{X}_p^T(k) \Delta \mathbf{w}(k) = \mathbf{e}_p(k) \quad (7)$$

where $\mathbf{d}_p(k)$ and $\mathbf{e}_p(k)$ are vectors with p elements, and $\mathbf{X}_p(k)$ is an $L \times p$ matrix defined by [14]:

$$\mathbf{d}_p(k) = [d(k), d(k-1), \dots, d(k-p+1)]^T \quad (8)$$

$$\mathbf{X}_p(k) = [x(k), x(k-1), \dots, x(k-p+1)] \quad (9)$$

$$\mathbf{e}_p(k) = \mathbf{d}_p(k) - \mathbf{X}_p^T(k) \mathbf{w}(k) \quad (10)$$

We can write the prior error equation vector as:-

$$\mathbf{e}_p(k) = [e(k), e(k-1), \dots, e(k-p+1)]^T \quad (11)$$

Since the number of simultaneous equations p in (5) is smaller than the number of unknowns L , these equations are indeterminate. Among the many solutions, the minimum norm solution can be derived from (7) by [13]:

$$\Delta \mathbf{w}(k) = \mathbf{X}_p(k) (\mathbf{X}_p^T(k) \mathbf{X}_p(k))^{-1} \mathbf{e}_p(k) \quad (12)$$

Substituting (12) into (4) gives the adjustment formula for the original projection algorithm that first time proposed by Ozeki/ Umeda [13]:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{X}_p(k) (\mathbf{X}_p^T(k) \mathbf{X}_p(k))^{-1} \mathbf{e}_p(k) \quad (13)$$

Table I. below shows the steps sequence of APA algorithm.

TABLE I: APA ALGORITHM

| |
|--|
| <p>1. Initialization:- Set $\mathbf{w}(0) = [0, 0, \dots, 0]^T$, and set the parameters L is the order of taps FIR filter, p is the projection order, and μ is the step size.</p> <p>Given the input vector:- $\mathbf{x}(k) = [x(k), x(k-1), \dots, x(k-L+1)]^T$, with dimension vector $L \times 1$, and $d(k)$ is the desired response at time k.</p> <p>$\mathbf{X}_p(k) = [x(k), x(k-1), \dots, x(k-p+1)]$</p> <p>$\mathbf{e}_p(k) = [e(k), e(k-1), \dots, e(k-p+1)]^T$, is the prior error equation vector with dimension $p \times 1$.</p> <p>$\mathbf{d}_p(k) = [d(k), d(k-1), \dots, d(k-p+1)]^T$, is the desired response vector with dimension $p \times 1$.</p> <p>For $k = 1, 2, \dots$ iterations Compute the following:- $\mathbf{e}_p(k) = \mathbf{d}_p(k) - \mathbf{X}_p^T(k) \mathbf{w}(k)$ $\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \mathbf{X}_p(k) (\mathbf{X}_p^T(k) \mathbf{X}_p(k))^{-1} \mathbf{e}_p(k)$</p> |
|--|

III. A NEW PROPOSED ALGORITHM

As described earlier this paper proposes a new algorithm which used adaptive step size instead of fixed one. This time varying step size is adjusted according to absolute mean value of the current and the previous estimation error's vector (11) as follows:

$$\mu(k+1) = \mu(k) (1 - \text{Abs}(\text{mean}(\mathbf{e}_p(k))) * \delta) \quad (14)$$

where $0 < \delta < 1$

$$\left. \begin{array}{l} \text{Then } \mu(k+1) = \mu_{MAX} \text{ if } \mu(k+1) > \mu_{MAX} \\ \text{Or } \mu(k+1) = \mu_{min} \text{ if } \mu(k+1) < \mu_{min} \\ \text{Otherwise } \mu(k+1) = \mu(k+1) \end{array} \right\} \quad (15)$$

The main idea of (14) is to make the step size take a large value at the beginning of the learning process, and then it decays gradually until it reaches a fixed selected value in the rest of the learning process.

The way in which $\mu(k+1)$ is changing depends on previous value of step size $\mu(k)$ and also on the absolute mean of estimation error vector $e_p(k)$. Number of error signal elements in $e_p(k)$ vector will depend upon the value of the projection order (p). The main reason of using error vector $e_p(k)$ in (14) is that to make use of gradually decreasing the error signal from large value to small one.

Furthermore, it is not required to recalculated again when the adaptive step size is adjusted and in turns the proposed formula in (14) is simple when it is implemented in hardware. As shown in equation (14) one can start with large step size, to enhance the convergence speed, and gradually reduce it to attain its minimum value, to achieve a low level of misadjustment. To achieve best performance the step size should decrease to the next, smaller step in smoothing transient and in an exponential manner.

Then substitute (14) into (13) and get:-

$$w(k+1) = w(k) + \mu(k)X_p(k) (X_p^T(k)X_p(k))^{-1}e_p(k) \quad (16)$$

To ensure stability, the variable step size $\mu(k+1)$ is constrained to the pre-determined maximum μ_{MAX} and minimum step size values μ_{min} , such that $\mu(k+1)$ is set to μ_{min} or μ_{MAX} when it falls below or above these lower and upper bounds, respectively [14]. The constant μ_{MAX} is normally selected near the point of instability of the conventional APA to provide the maximum possible convergence speed. Given that the initial value of step size $\mu(1)$ (i.e. when k=1) equals to μ_{MAX} . The value of μ_{min} is chosen as a compromise between the desired level of steady state misadjustment and the required tracking capabilities of the algorithm. The parameter δ controls the convergence time as well as the level of misadjustment of the algorithm at a steady state. The profile change of time varying step size of (14) will be an exponential form as will be seen later in the next section.

Table II shows main step sequences for the proposed algorithm (AMASSAPA).

IV. SIMULATION RESULTS

The proposed algorithm and conventional APA have been tested, by applying them to a system identification model as shown in Fig. 2.

As shown an adaptive filter is used in the system identification model. The adaptive filter and unknown system are FIR and both have the same number of taps (L= 36). The unknown system has 36 randomly selected taps. The goal of the adaptive processing is to estimate unknown system parameters (weight coefficients W) by optimizing the adaptive filter parameters (weight coefficients W[^]) iteratively using APA and new proposed AMASSAPA algorithm as shown in Tables I and II.

An additive noise n(k) with variance 0.5 is added to the output of an unknown system model in order to disturbance y[^](k) and make the estimation process more difficult.

The input x(k) and disturbance n(k) signals are uncorrelated with each other over time. The input signal x(k) (colored signal) is obtained by filtering a white Gaussian random noise x(k) (zero mean, unit variance) through a 1st order filter (first-order autoregressive system), $A(z) = 1/(1 - 0.9z^{-1})$, where z^{-1} is the sample delay operator, the value of $0.9z^{-1}$ is chosen to generate a highly colored Gaussian signal.

TABLE II: A NEW PROPOSED AMASSAPA ALGORITHM

| |
|---|
| <p>1. Initialization:- Set $w(0) = [0, 0, \dots, 0]^T$, and set the parameters L is the order of taps FIR filter, p is the projection order, δ, μ_{min} and μ_{MAX} for adjusting step size. Given the input vector:- $x(k) = [x(k), x(k-1), \dots, x(k-L+1)]^T$, with dimension vector (LXp), and d(k) is the desired response at time k. $X_p(k) = [x(k), x(k-1), \dots, x(k-p+1)]$ $e_p(k) = [e(k), e(k-1), \dots, e(k-p+1)]^T$, is the prior error equation vector with dimension px1. $d_p(k) = [d(k), d(k-1), \dots, d(k-p+1)]^T$, is the desired response vector with dimension px1.</p> <p>2. For k= 1,2,..... Iterations Compute the following:- $e_p(k) = d_p(k) - X_p^T(k)w(k)$ $\mu(1) = \mu_{MAX}$, $\mu(k+1) = \mu(k)(1 - Abs(mean(e_p(k))) * \delta)$</p> <p>$\mu(k+1) = \mu_{MAX}$ if $\mu(k+1) > \mu_{MAX}$ Or $\mu(k+1) = \mu_{min}$ if $\mu(k+1) < \mu_{min}$ Otherwise $\mu(k+1) = \mu(k+1)$</p> <p>$w(k+1) = w(k) + \mu(k)X_p(k)(X_p^T(k)X_p(k))^{-1}e_p(k)$</p> |
|---|

The step size μ used for APA is equal to 0.08, and the parameters for AMASSAPA are as following:-, $\delta = 1e - 3$, $\mu_{MAX} = 0.08$, and $\mu_{min} = 1e - 6$. All the values of these parameters were chosen to achieve better performance in terms of fast convergence time and low level misadjustment for both algorithms in order to make a fairly comparison between these algorithms.

Fig. 3 shows the squared estimation error between unknown system weight coefficients (W) and the corresponding FIR weight coefficients (W[^]) i.e. (W-W[^]) for APA and AMASSAPA algorithms when projection order (p) is equal to 2. In this case, the number of elements in $e_p(k)$ vector that used for adjusting step size is equal to two (i.e. $e_p(k) = [e(k), e(k-1)]$).

As shown in Fig.3, the AMASSAPA has fast convergence time (50 iterations) than APA (100 iterations).

Fig. 4 shows the performance of both algorithms when the projection order was 4. In this case the number of elements in $e_p(k)$ vector that used for adjusting step size is equal to four (i.e. $e_p(k) = [e(k), e(k-1), e(k-2), e(k-3)]$). As shown in Fig.4, proposed algorithm AMASSAPA has fast

convergence time (100 iterations) compared with APA (400 iterations). Fig.5 shows the performance of both algorithms when the projection order was increased to 8.

In this case the number of elements in $e_p(k)$ vector that used for adjusting step size is equal to eight (i.e. $e_p(k) = [e(k), e(k-1), e(k-2), e(k-3), \dots, e(k-7)]$).

It is clear from Fig. 5, that the proposed algorithm AMASSAPA also has fast convergence time (200 iterations) compared with APA (1000 iterations).

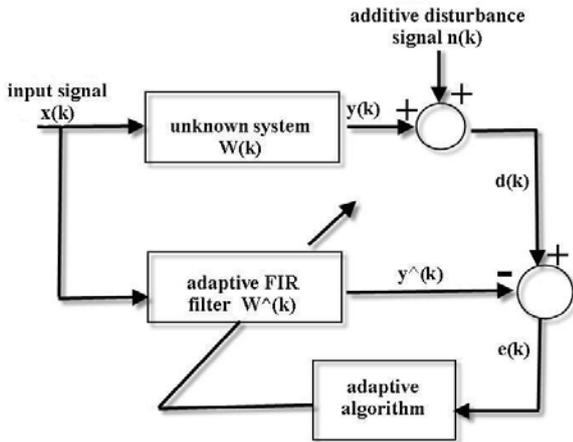


Fig. 2. System identification structure

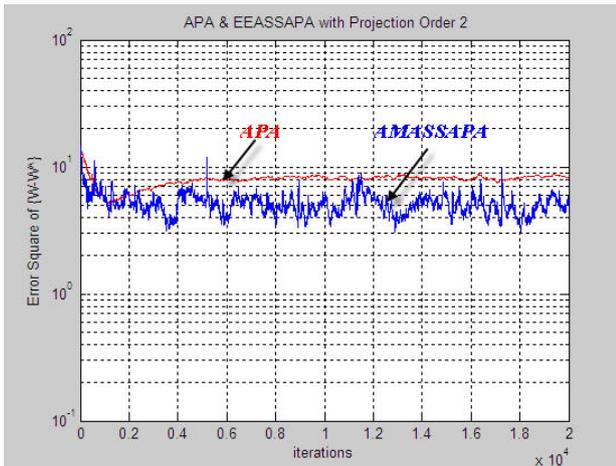


Fig. 3. Squared estimation error ($W-W^{\wedge}$) for APA and AMASSAPA algorithms when projection order (P) is equal to 2.

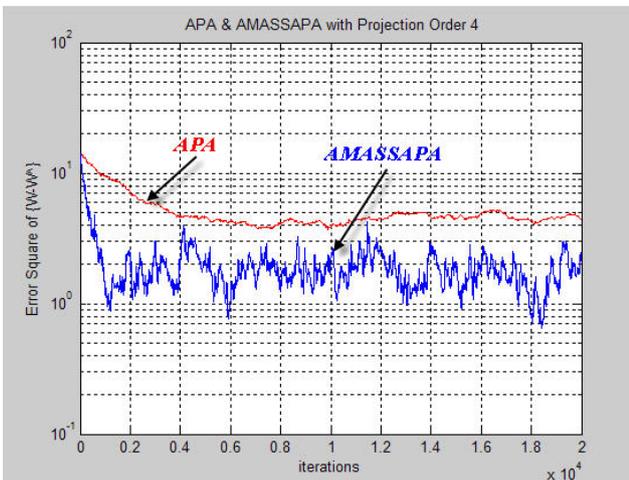


Fig. 4. Squared estimation error ($W-W^{\wedge}$) for APA and AMASSAPA algorithms when projection order (P) is equal to 4.

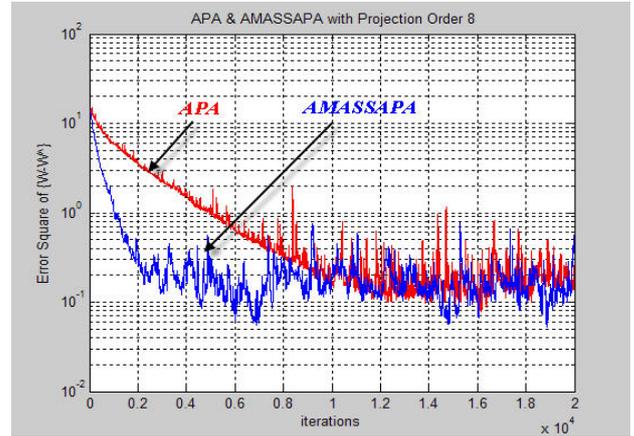


Fig. 5. Squared estimation error ($W-W^{\wedge}$) for APA and AMASSAPA algorithms when projection order (p) is equal to 8.

Fig. 6 shows the performance of both algorithms when the projection order was increased to sixteen. In this case, the number of elements in $e_p(k)$ vector that used for adjusting step size is equal to 16 (i.e. $e_p(k) = [e(k), e(k-1), e(k-2), e(k-3), \dots, e(k-15)]$).

As shown in Fig.6, the proposed algorithm AMASSAPA has also fast convergence time (600 iterations) compared with APA which has convergence time equal to 1100 iterations.

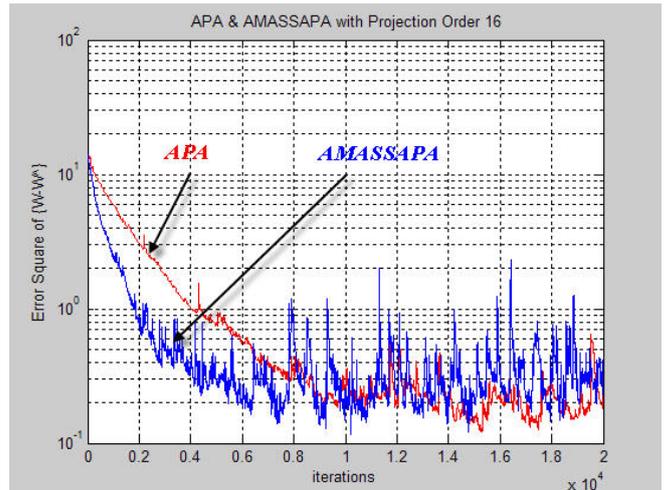


Fig. 6. Squared estimation error ($W-W^{\wedge}$) for APA and AMASSAPA algorithms when projection order (p) is equal to 16.

It is clear from these figures that the performance of AMASSAPA is improved very well compared with APA when projection order value is increased. The reason of that is the number of error vector e used for adaptation is increased and the algorithm applies to update directions that are orthogonal to the last p input vectors and thus in turns allows decorrelation of an input process, speeding up the convergence.

Finally Fig. 7 shows the adaptive step size profile of $\mu(k+1)$ using AMASSAPA for a different projection orders. As shown in this figure the profile has an exponential form, and its curvatures are enhanced when a projection order is increased.

V. CONCLUSIONS

In this paper possible performance enhancement of colored adaptive filtering was achieved using new proposed adjusted

step size affine projection algorithm (AMASSAPA). The performance of the proposed algorithm was illustrated by simulations of a system identification model. Through simulation results, the proposed algorithms show fast and low level of missadjustment compared with fixed step size affine projection algorithm (APA). The profile of adaptive step size is enhanced when a higher projection order was used because more prior error signals were included, which in turns improve the adaptation process.

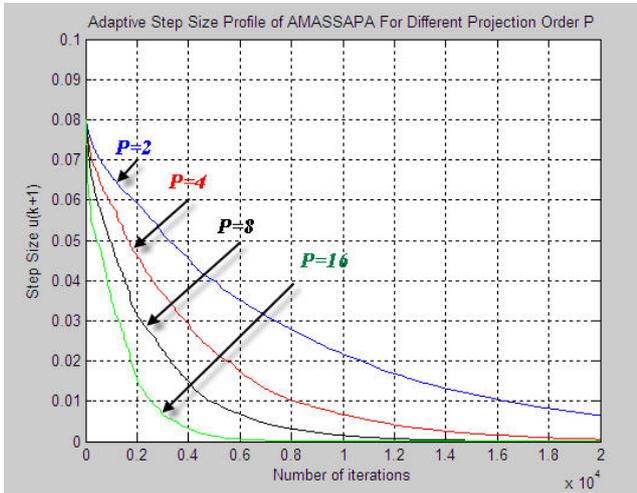


Fig. 7. Adaptive Step Size Profile of AMASSAPA for Different Projection Order P

REFERENCES

[1] T. Dai, A. Adler, and B. Shahrava, "Variable Step-Size Affine Projection Algorithm with a Weighted and Regularized Projection Matrix," *International Journal of Computer, Information, and Systems Science, and Engineering*, vol. 2, no. 3, Summer 2008.

[2] M. Rupp, "A Family of Adaptive Filter Algorithms with Decorrelating Properties," *IEEE Transactions on Signal Processing*, vol. 46, no. 3, March, 1998.

[3] M. S. E. Abadi, V. Mehrdad, A. Gholipour, and M. Noroozi, "A Family of Variable Step-Size Affine Projection Adaptive Filtering Algorithms," *International Journal of Computer and Electrical Engineering*, vol. 2, no. 3, June, 2010.

[4] K. Y. Hwang and W. J. Song, "An Affine Projection Adaptive Filtering Algorithm with Selective Regressors," In *IEEE Transactions on Circuits and Systems—II: Express Briefs*, vol. 54, no. 1, January, 2007.

[5] H. C. Shin and A. H. Sayed, "Mean-Square Performance of a Family of Affine Projection Algorithms," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, January, 2004.

[6] C. Paleologu, J. Benesty, and S. Ciochin, "A Variable Step-Size Affine Projection Algorithm Designed for Acoustic Echo Cancellation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 16, no. 8, November, 2008.

[7] F. Albu and C. Paleologu, "The Variable Step-Size Gauss-Seidel Pseudo Affine Projection Algorithm," *World Academy of Science, Engineering and Technology*, vol. 49, 2009.

[8] A. Trivedii and D. K. Mehra, "Variable Step Size Affine Projection Algorithm for Adaptive Multiuser DS-CDMA MMSE Receiver," *Wireless Personal Communications*, vol. 42, pp. 107–130.

[9] L. Liu, M. Fukumoto, S. Saiki, and S. Zhang, "A Variable Step-Size Proportionate Affine Projection Algorithm for Identification of Sparse Impulse Response," *Research Article, Hindawi Publishing Corporation EURASIP Journal on Advances in Signal Processing*, vol. 2009, Article ID 150914, August 2009.

[10] M. Çabuk, "Adaptive Step Size and Exponentially Weighted Affine Projection Algorithms," *Master of Science Thesis in Electrical and Electronics Engineering Boğaziçi University*, 2002.

[11] K. Sakhnov, "An Improved Proportionate Affine Projection Algorithm for Network Echo Cancellation," *Department of Telecommunication Engineering, Czech Technical University in Prague*, 2002.

[12] B. Widrow and S. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Inc. 1985.

[13] Y. Kaneda, M. Tanaka, and J. Kajima, "An adaptive algorithm with fast convergence for multi input sound control," In Newport Beach, CA, USA, 1995.

[14] R. H. Kang, E. W. Johnstone, "A variable step size LMS algorithm," *IEEE Trans. On Signal Processing*, vol.40, no.7, pp.1633-1642, 1992.



Thamer M. Jamel was born in Baghdad , Iraq in 1961. He graduated from the University of Technology with a Bachelor's degree in electronics engineering in 1983. He received a Master's degree in digital communications engineering from the University of Technology in 1990 and a Doctoral degree in communication engineering from the University of Technology in 1997. Currently, he is associate professor in the communication engineering branch, Electrical engineering department at University of Technology, Baghdad, Iraq. His research interests include adaptive signal processing , Neural Networks , DSP microprocessors , FPGA , and Modern digital communications systems.