A THEORETICAL MODEL FOR THE DEFICIENT ORDER PSEUDO-AFFINE PROJECTION ALGORITHM

Sérgio J. M. de Almeida[1], Márcio H. Costa[2], and José C. M. Bermudez[2]

[1]Centro Politécnico, Universidade Católica de Pelotas, Pelotas-RS, Brazil
[2]Departamento de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis-SC, Brazil
E-mail: smelo@ucpel.tche.br; costa@eel.ufsc.br; j.bermudez@ieee.org

ABSTRACT

This paper presents a statistical analysis of the Pseudo-Affine Projection (PAP) adaptive algorithm when the order of the PAP algorithm is smaller than the order of the autoregressive (AR) input process. Deterministic recursive equations are derived for the mean weight and mean-square error behavior. Monte Carlo simulations show good agreement with the theoretical predictions in steady-state and during transient. These results are of special interest in practical applications where the computational complexity prevents implementation of the sufficient order PAP algorithm for high order AR inputs.

1. INTRODUCTION

The Affine Projection (AP) adaptive algorithm [1] is nowadays recognized as an attractive alternative to speed up convergence of gradient-based algorithms such as LMS and NLMS. The AP algorithm applies weight updates in directions that are orthogonal to the last P input vectors. This decorrelates the input signal and speeds up convergence [2]. The improved transient performance comes at the cost of increased computational complexity and steady-state misadjustment.

A simplified version of the AP algorithm has been proposed in [3]. The pseudo-AP (PAP) adaptive algorithm replaces the input signal with its autoregressive (AR) prediction. In this case, the error vector in the conventional AP weight update turns to a scalar for unit step size, resulting in a lower computational cost compared to the conventional AP algorithm. The PAP algorithm uses this simplified AP weight update but with step size (α) less than one. The PAP presents the same behavior as the AP adaptive algorithm only when the input signal is truly AR and α=1. Using α<1 provides a tradeoff between steady-state misadjustment and convergence speed.

An analytical model has been derived in [4] for the PAP with $\alpha \le 1$, AR inputs of known order and an adaptive filter with sufficient length. In [5], this model has been extended to the deficient length case (when the length of the plant is underestimated). Despite these preliminary results, the study of the PAP algorithm behavior still represents a challenge.

One important unsolved practical issue is the behavior of PAP with deficient order, when the chosen value of P is insufficient for a proper modeling of the input process. Such situation is common in practical applications. The computational complexity of an AP algorithm with order H is of order NK^2 for N adaptive weights. Thus, implementation costs may become prohibitive for large values of K. In such case, simulation studies show steady-state performance losses that cannot be predicted by the available models. A recent paper [6] presented analytical models for the mean weight and mean square error behaviors of the deficient AP algorithm for unity step-size and autoregressive signals. However, the developed models cannot be directly generalized to the PAP algorithm. This work is a study of the deficient order PAP algorithm behavior in system identification.

The paper is organized as follows. Section 2 introduces the input signal model and the notation used. Section 3 formulates the deficient order AP weight update equation. Section 4 derives the analytical model for the algorithm behavior. Section 5 presents Monte Carlo simulations to validate the theoretical model. Section 6 concludes the work.

2. THE INPUT SIGNAL MODEL

The adaptive system attempts to estimate a desired signal $d(n)$ that can be modeled by

$$d(n) = \mathbf{w}^{\mathbf{o}T}\mathbf{u}(n) + r(n) \quad (1)$$

where the N-length vector $\mathbf{w^o}=[w^o_0\ w^o_1\ w^o_2\ \dots\ w^o_{N-1}]^T$ models the impulse response of the unknown system (plant) and $\mathbf{u}(n)=[u(n)\ u(n-1)\ \dots\ u(n-N+1)]^T$ is the input regressor, with autocorrelation matrix $\mathbf{R}_u=E\{\mathbf{u}(n)\mathbf{u}^T(n)\}$. The input signal $u(n)$ is assumed to be a stationary AR process of order H. Such process can model a variety of input signals found in practical applications. Thus,

$$\mathbf{u}(n) = \sum_{i=1}^{H} a_i \mathbf{u}(n-i) + \mathbf{z}(n) \quad (2)$$

where $\mathbf{u}(n-k)=[u(n-k)\ u(n-k-1)\ \dots\ u(n-k-N+1)]^T$ and $\mathbf{z}(n)=[z(n)\ z(n-1)\ \dots\ z(n-N+1)]^T$ is a vector with samples from a stationary white Gaussian process with variance σ_z^2.

3. THE DEFICIENT ORDER PAP ALGORITHM

The weight update equation for the PAP algorithm can be written as [4], [7]:

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha \frac{\mathbf{\Phi}(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)} e(n) \quad (3)$$

where $\mathbf{w}(n)=[w_0(n)\ w_1(n)\ \dots\ w_{N-1}(n)]^T$ is the adaptive weight vector and the error signal $e(n)$ is given by

$$e(n) = \mathbf{w}^{oT}\mathbf{u}(n) + r(n) - \mathbf{w}^T(n)\mathbf{u}(n) \tag{4}$$

Vector $\mathbf{\Phi}(n)=[\phi_0, \phi_1, \ldots, \phi_{N-1}]^T$ defines the update direction, and is given by:

$$\mathbf{\Phi}(n) = \mathbf{u}(n) - \mathbf{U}(n)\hat{\mathbf{a}}(n) \tag{5}$$

where $\mathbf{U}(n)=[\mathbf{u}(n\text{-}1)\ \mathbf{u}(n\text{-}2)\ \ldots\ \mathbf{u}(n\text{-}P)]$. The PAP order is ($P$+1), where P is the number of past input vectors used to determine $\mathbf{\Phi}(n)$. The least squares (LS) estimate $\hat{\mathbf{a}}(n)$ of the AR coefficient vector $\mathbf{a}$ is given by:

$$\hat{\mathbf{a}}(n) = [\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)\mathbf{u}(n) \tag{6}$$

where $\hat{\mathbf{a}}(n)=[\hat{a}_1(n)\ \hat{a}_2(n)\ \ldots\ \hat{a}_P(n)]^T$ and $\mathbf{U}^T(n)\mathbf{U}(n)$ is assumed to have rank P.

To study the effect of a deficient order in the PAP algorithm we assume $P<H$ and express (2) as

$$\mathbf{u}(n) = \mathbf{U}(n)\mathbf{b} + \bar{\mathbf{U}}(n)\mathbf{c} + \mathbf{z}(n) \tag{7}$$

where $\bar{\mathbf{U}}(n)=[\mathbf{u}(n\text{-}P\text{-}1)\ \mathbf{u}(n\text{-}P\text{-}2)\ \ldots\ \mathbf{u}(n\text{-}H)]$ contains the (H-P) past input vectors not included in $\mathbf{U}(n)$, $\mathbf{b}=[a_1\ a_2\ \ldots\ a_P]^T$ and $\mathbf{c}=[a_{P+1}\ a_{P+2}\ \ldots\ a_H]^T$. Using (7) in (6) results in

$$\hat{\mathbf{a}}(n) = \mathbf{b} + [\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)\left[\bar{\mathbf{U}}(n)\mathbf{c} + \mathbf{z}(n)\right] \tag{8}$$

Equation (8) shows that the insufficient order LS estimate of the AR coefficients is a biased estimate of the real first P weights in $\mathbf{b}$. Thus, $\mathbf{\Phi}(n)$ obtained from (5) can no longer be modeled as a white vector sequence as happened in the sufficient order PAP [4].

4. THEORETICAL ANALYSIS

The following statistical assumptions are used in the analysis. A detailed discussion of these assumptions can be found in [4] and [7]:

- A1: The number of adaptive filter is large enough so that $N>>P$.
- A2: The statistical dependence between z(n) and $\mathbf{U}(n)$ can be neglected for $N>>P$.
- A3: Vector $\mathbf{\Phi}(n)$ is orthogonal to the columns of $\mathbf{U}(n)$.
- A4: Vectors $\mathbf{\Phi}(n)$ and $\mathbf{w}(n)$ are statistically independent.

4.1 Mean Weight Behavior

Defining the weight-error vector as $\mathbf{v}(n)=\mathbf{w}(n)-\mathbf{w}^o$, equation (3) can be written as

$$\mathbf{v}(n+1) = \mathbf{v}(n) + \alpha\frac{\mathbf{\Phi}(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)}e(n) \tag{9}$$

Using (4) and (5) in (9) yields

$$\mathbf{v}(n+1) = \mathbf{v}(n) - \frac{\alpha}{\gamma}\frac{\mathbf{\Phi}(n)\mathbf{\Phi}^T(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)}\mathbf{v}(n) + \alpha\frac{\mathbf{\Phi}(n)r_a(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)} \tag{10}$$

where

$$r_a(n) = r(n) - \frac{\alpha}{\gamma}\hat{\mathbf{a}}^T(n)\mathbf{r}(n-1) \tag{11}$$

is the filtered noise sequence and $\gamma=1-(1-\alpha)\Sigma^P_{i=1}\hat{a}_i$ [4]. It was assumed that $\hat{\mathbf{a}} \cong E\{\hat{\mathbf{a}}(n)\} = \mathbf{b} + \mathbf{\Sigma}\mathbf{c}$, where

$$\mathbf{\Sigma} = E\{[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)\bar{\mathbf{U}}(n)\} \tag{12}$$

Taking the expected value of (10) and using assumption A4 yields

$$\begin{aligned} E\{\mathbf{v}(n+1)\} &= E\{\mathbf{v}(n)\} - \frac{\alpha}{\gamma}E\left\{\frac{\mathbf{\Phi}(n)\mathbf{\Phi}^T(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)}\right\}E\{\mathbf{v}(n)\} \\ &+ \alpha E\left\{\frac{\mathbf{\Phi}(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)}r_a(n)\right\} \end{aligned} \tag{13}$$

Using A1, the first expected value in (13) can be approximated as in [7], but considering the nonwhite characteristic of $\mathbf{\Phi}(n)$ [8]. Hence,

$$E\left\{\frac{\mathbf{\Phi}(n)\mathbf{\Phi}^T(n)}{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)}\right\} \cong \frac{1}{(G-2)\sigma_\Phi^2}\mathbf{R}_\Phi \tag{14}$$

where $G=N\text{-}P$ and σ_Φ^2 is the variance of the elements of $\mathbf{\Phi}(n)$ and $\mathbf{R}_\Phi=E\{\mathbf{\Phi}^T(n)\mathbf{\Phi}(n)\}$. Since $E\{\mathbf{\Phi}(n)r_a(n)\}=0$ ($r(n)$ is zero-mean and independent of any other signal), using (14) in (13) yields

$$E\{\mathbf{v}(n+1)\} = \left[\mathbf{I} - \frac{\alpha}{\gamma}\frac{\mathbf{R}_\Phi}{(G-2)\sigma_\Phi^2}\right]E\{\mathbf{v}(n)\} \tag{15}$$

Note that, differently from the sufficient order PAP case (P=H) [4], matrix $\mathbf{R}_\Phi$ in (15) is not diagonal.

Analyzing (15) for $n\to\infty$ it is easy to verify that the deficient order PAP algorithm produces an unbiased solution since

$$\lim_{n\to\infty} E\{\mathbf{v}(n)\} = \mathbf{0} \tag{16}$$

4.2 Matrix $\mathbf{R}_\Phi$

Substituting (8) in (5), post-multiplying the resulting expression by its transpose and taking the expected value, the autocorrelation matrix of $\mathbf{\Phi}(n)$ can be written as

$$\begin{aligned} \mathbf{R}_\Phi &= E\left\{\mathbf{z}_\perp(n)\mathbf{z}_\perp^T(n)\right\} \\ &+ E\left\{\mathbf{z}_\perp(n)\mathbf{c}^T\bar{\mathbf{U}}^T(n)\mathbf{P}_\perp^T(n)\right\} \\ &+ E\left\{\mathbf{P}_\perp(n)\bar{\mathbf{U}}(n)\mathbf{c}\mathbf{z}_\perp^T(n)\right\} \\ &+ E\left\{\mathbf{P}_\perp(n)\bar{\mathbf{U}}(n)\mathbf{c}\mathbf{c}^T\bar{\mathbf{U}}^T(n)\mathbf{P}_\perp^T(n)\right\} \end{aligned} \tag{17}$$

where $\mathbf{P}_U(n)=\mathbf{U}(n)[\mathbf{U}^T(n)\mathbf{U}(n)]^{-1}\mathbf{U}^T(n)$ is the projection matrix onto the subspace spanned by the columns of the $\mathbf{U}(n)$ and $\mathbf{P}_\perp(n)=(\mathbf{I}-\mathbf{P}_U(n))$ is the projection matrix onto its orthogonal complement.

The first expectation in (17) was already solved in [7]:

$$E\left\{\mathbf{z}_\perp(n)\mathbf{z}_\perp^T(n)\right\} = \frac{N-P}{N}\sigma_z^2\mathbf{I} \tag{18}$$

Extensive simulations have shown that the second and the third terms in (17) are small when compared to the first and fourth terms. Neglecting these terms in (17) yields

$$\mathbf{R}_\Phi \cong \frac{N-P}{N}\sigma_z^2\mathbf{I} + \mathbf{\Upsilon} \tag{19}$$

where $\mathbf{\Upsilon}=E\{\mathbf{P}_\perp(n)\bar{\mathbf{U}}(n)\mathbf{c}\mathbf{c}^T\bar{\mathbf{U}}(n)\mathbf{P}_\perp^T(n)\}$ depends only on the input statistics.

4.3 Mean Square Error

Squaring (4), using the weight-error vector definition and taking the expected value yields

$$E\{e^2(n)\} = E\{r_a^2(n)\} + tr[\mathbf{R}_\Phi\mathbf{K}(n)] \tag{20}$$

where $\mathbf{K}(n)=E\{\mathbf{v}(n)\mathbf{v}^T(n)\}$ is the weight-error correlation matrix. $E\{r_a^2(n)\}$ can be obtained using (8) in (11) and taking the expectation. Thus,

$$E\{r_a^2(n)\} \cong \left[1+\frac{\alpha^2}{\gamma^2}\left(\mathbf{b}^T\mathbf{b}+2\mathbf{b}^T\boldsymbol{\Sigma}\mathbf{c}+\mathbf{c}^T\boldsymbol{\Pi}\mathbf{c}\right)\right]\sigma_r^2 \quad (21)$$

where

$$\boldsymbol{\Pi} = E\{\bar{\mathbf{U}}^T(n)\mathbf{U}(n)[\mathbf{U}^T(n)\mathbf{U}(n)]^{-2}\mathbf{U}^T(n)\bar{\mathbf{U}}(n)\} \quad (22)$$

Using (21) in (20) results in

$$E\{e^2(n)\} = \left[1+\frac{\alpha^2}{\gamma^2}\left(\mathbf{b}^T\mathbf{b}+2\mathbf{b}^T\boldsymbol{\Sigma}\mathbf{c}+\mathbf{c}^T\boldsymbol{\Pi}\mathbf{c}\right)\right]\sigma_r^2 + tr[\mathbf{R}_\Phi\mathbf{K}(n)] \quad (23)$$

To evaluate the mean square error it is still necessary to model the evolution of the weight-error correlation matrix. This will be done in the next section.

4.4 Weight-Error Correlation Matrix

A recursive expression for $\mathbf{K}(n)$ can be obtained as done in [4] and [7]. Post-multiplying (10) by its transpose, taking the expected value and applying assumptions A1 to A4 yields (for details see [4]):

$$\begin{aligned}\mathbf{K}(n+1) &= \mathbf{K}(n) - \frac{\alpha}{\gamma\sigma_\Phi^2(G-2)}\left[\mathbf{K}(n)\mathbf{R}_\Phi + \mathbf{R}_\Phi\mathbf{K}(n)\right] \\ &+ \frac{\alpha^2}{\gamma^2}E\left\{\frac{\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)}{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^2}\right\} \\ &+ \alpha^2 E\left\{\frac{r_a^2(n)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)}{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^2}\right\}\end{aligned} \quad (24)$$

The first expectation in (24) can be approximated as [7]

$$\begin{aligned}&E\left\{\frac{\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\mathbf{v}(n)\mathbf{v}^T(n)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)}{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^2}\right\} \\ &\cong E\left\{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^{-2}\right\} \\ &\times E\left\{tr\left\{\mathbf{v}(n)\mathbf{v}^T(n)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right\}\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right\}\end{aligned} \quad (25)$$

where

$$E\left\{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^{-2}\right\} \cong \frac{1}{\sigma_\Phi^4 G(G+2)} \quad (26)$$

Differently from the results in [7] $E\{\Phi_i(n)\Phi_j(n)\}\neq 0$ for $i\neq j$. We approximate the expectation in (25) by

$$\begin{aligned}&E\left\{tr\left\{\mathbf{v}(n)\mathbf{v}^T(n)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right\}\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right\} \\ &\cong tr\left\{\mathbf{K}(n)\mathbf{R}_\Phi\right\}\mathbf{R}_\Phi\end{aligned} \quad (27)$$

The last expected value in (24) is approximated by [7]

$$\begin{aligned}&E\left\{\frac{r_a^2(n)\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)}{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^2}\right\} \\ &\cong E\left\{r_a^2(n)\right\}E\left\{\left[\boldsymbol{\Phi}^T(n)\boldsymbol{\Phi}(n)\right]^{-2}\right\}E\left\{\boldsymbol{\Phi}(n)\boldsymbol{\Phi}^T(n)\right\}\end{aligned} \quad (28)$$

The solution to the second and third expected values of (28) can be found in [7] for $\boldsymbol{\Phi}(n)$ white. However, following the results in [8], through assumption A1, we can use the same approximation for $\boldsymbol{\Phi}(n)$ correlated. Proceeding as in [4] and using A1, it can be shown that (24) becomes

$$\begin{aligned}\mathbf{K}(n+1) &= \mathbf{K}(n) - \frac{\alpha}{\gamma\sigma_\Phi^2(G-2)}\left[\mathbf{K}(n)\mathbf{R}_\Phi + \mathbf{R}_\Phi\mathbf{K}(n)\right] \\ &+ \frac{\alpha^2}{\gamma^2\sigma_\Phi^4 G(G+2)}tr\left\{\mathbf{K}(n)\mathbf{R}_\Phi\right\}\mathbf{R}_\Phi \\ &+ \frac{\alpha^2\left[1+\frac{\alpha^2}{\gamma^2}\left(\mathbf{b}^T\mathbf{b}+2\mathbf{b}^T\boldsymbol{\Sigma}\mathbf{c}+\mathbf{c}^T\Pi\mathbf{c}\right)\right]\sigma_r^2}{\sigma_\Phi^4(G-2)(G-4)}\mathbf{R}_\Phi\end{aligned} \quad (29)$$

which completes the model for the mean-square error.

4.5 Steady-State Mean Square Error

The steady-state mean square error behavior can be determined from (23) by assuming convergence as $n\to\infty$. Using $\mathbf{K}(n+1)=\mathbf{K}(n)=\mathbf{K}_\infty$ in (29) as $n\to\infty$, the trace of (29) becomes

$$tr\left\{\mathbf{K}_\infty\mathbf{R}_\Phi\right\} = \frac{\alpha\gamma^2 NG(G+2)\left[1+\frac{\alpha^2}{\gamma^2}\left(\mathbf{b}^T\mathbf{b}+2\mathbf{b}^T\boldsymbol{\Sigma}\mathbf{c}+\mathbf{c}^T\Pi\mathbf{c}\right)\right]\sigma_r^2}{2\gamma G(G+2)(G-4)-\alpha N(G-2)(G-4)} \quad (30)$$

where $tr[\mathbf{R}_\Phi]=N\sigma_\Phi^2$. Using (30) in (23) we obtain

$$\lim_{n\to\infty}E\{e^2(n)\} = \left[1+\frac{\alpha^2}{\gamma^2}\left(\mathbf{b}^T\mathbf{b}+2\mathbf{b}^T\boldsymbol{\Sigma}\mathbf{c}+\mathbf{c}^T\boldsymbol{\Pi}\mathbf{c}\right)\right]\sigma_r^2 \times \frac{\left[2\gamma G(G+2)-\alpha N(G-2)\right](G-4)+\alpha\gamma^2 NG(G+2)}{\left[2\gamma G(G+2)-\alpha N(G-2)\right](G-4)} \quad (31)$$

4.6 Model's Restrictions

Through exhaustive simulations, it has been verified that the derived model provides good predictions of the algorithm behavior in the steady-state phase of adaptation. In general, the model's accuracy during the initial transient phase decreases as $|\mathbf{v}(0)|^2$ and σ_u^2/σ_z^2 increases. This occurs due to the recursive procedure used to obtain Equation (12d) in [2], which assumes the algorithm is always adapting (even before n=0). To overcome such limitation is a very difficult challenge and remains an open problem in the study of the AP and PAP algorithms.

5. SIMULATIONS

This section presents simulations to verify the accuracy and limitations of the analytical models given by equations (15), (23) and (29). In all cases, the statistics of the input signal (matrices $\boldsymbol{\Upsilon}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Pi}$) have been numerically estimated from the input process. The plant is obtained from a 64-tap Hanning window with unit norm; the weights were initialized with zero values; $\sigma_r^2 = 10^{-6}$; $\sigma_u^2 = 1$ and 150 runs. The input signal was obtained through an eighth order AR model (H=10) with coefficients $\mathbf{a}$=[-0.9 0.7 -0.6 0.5 -0.45 0.35 -0.3 0.25 -0.2 0.1] and $\sigma_z^2 = 0.538$. Fig. 1 and 2 show excess mean square error ($E\{e^2(n)\}- \sigma_r^2$) simulations and theoretical predictions for two deficient cases (P=2 and P=4) for α=0.4 and α=0.9. Figs. 3 to 6 show the mean weight behavior for coefficients 20, 30 and 40 for different sets of parameters. Horizontal lines represent the true values of the weights to be identified. Analysis of the obtained curves in steady-state agrees with the theoretical result predicted by (16) and (31) for different values of P.

6. CONCLUSIONS

This paper presented analytical models for predicting the stochastic behavior of the deficient order pseudo-Affine projection algorithm. Deterministic recursive equations were derived for the mean weight and mean square error. Simulation results have shown good agreement with theoretical predictions in steady-state. During the transient phase, good matches between theoretical models and simulations are obtained as the used theoretical assumptions are assured.

ACKNOWLEDGMENTS

The authors would like to thank CNPq (Brazilian Ministry of Science and Technology) for the financial support (grants 303803/2009-6 and 559418/2008-6).

REFERENCES

[1] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using orthogonal projection to an affine subspace and its properties," *Electron. Commun. Jpn.*, vol 67-A, no.5, pp. 19-27, Feb. 1984.

[2] M. Rupp, "A family of filter algorithms with decorrelating properties," *IEEE Transactions on Signal Processing*, vol. 46, no. 3, pp. 771–775, Mar. 1998.

[3] F. Bouteille, P. Scalart, and M. Corazza, "Pseudo affine projection algorithm new solution for adaptive identification," in *Proc. EUROSPEECH*, 1999, pp. 427–430.

[4] S.J.M. Almeida, J.C.M. Bermudez, and N.J. Bershad, "A stochastic Model for a Pseudo Affine Projection algorithm," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 107–118, Jan. 2009.

[5] S.J.M. Almeida, , M.H. Costa and , J.C.M. Bermudez, "A stochastic model for the deficient length pseudo Affine projection adaptive algorithm", in *Proc. of EUSIPCO* 2009, pp. 1715-1719.

[6] S.J.M. Almeida, , M.H. Costa and , J.C.M. Bermudez, "A stochastic model for the deficient order affine projection algorithm", in *Proc. of ISPAA* 2010, vol. 1, pp. 554-557.

[7] S.J.M. Almeida, J.C.M. Bermudez, N.J. Bershad, and M.H. Costa, "A statistical analysis of the Affine Projection algorithm for unity step size and autoregressive inputs," *IEEE Transactions on Circuits and Systems – I*, vol. 52, no. 7, pp. 1394–1405, Jul. 2005.

[8] M.H. Costa, J.C.M. Bermudez, "An improved model for the normalized LMS algorithm with Gaussian inputs and large number of coefficients," in *Proc. of ICASSP* 2002, vol. 2, pp. 1385-1388.

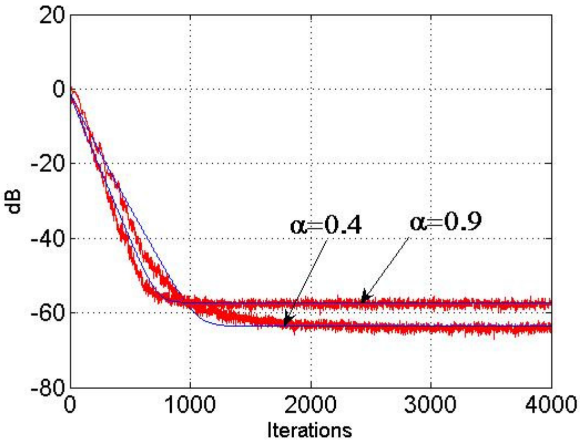

Figure 1 – Excess mean square error for P=2. Simulations (ragged - red) and model (continuous - blue).

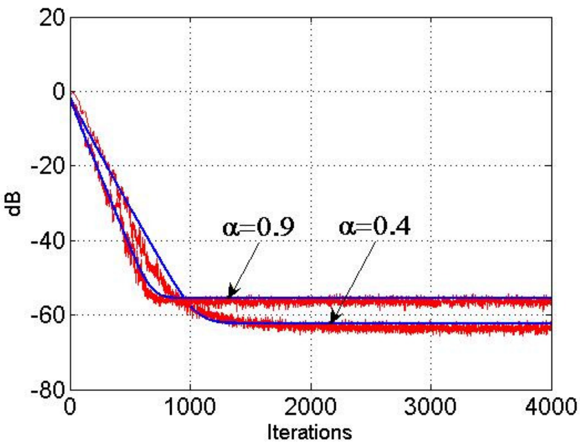

Figure 2 – Excess mean square error for P=4. Simulations (ragged - red) and model (continuous - blue).

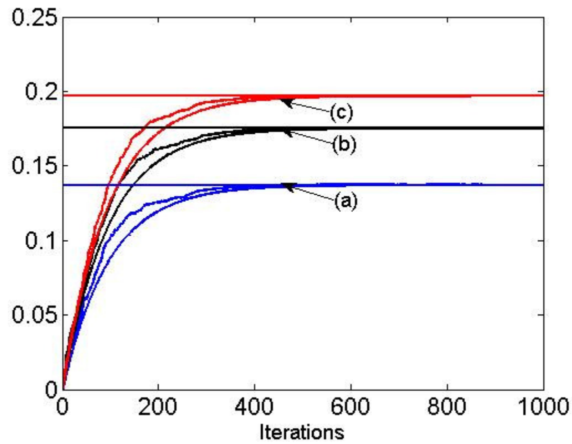

Figure 3 – Mean weight behavior for P=2 and α=0.4. Simulations (ragged) and model (continuous). (a) coefficient 20; (b) coefficient 40; (c) coefficient 30. Horizontal lines represent the true values of the plant coefficients.

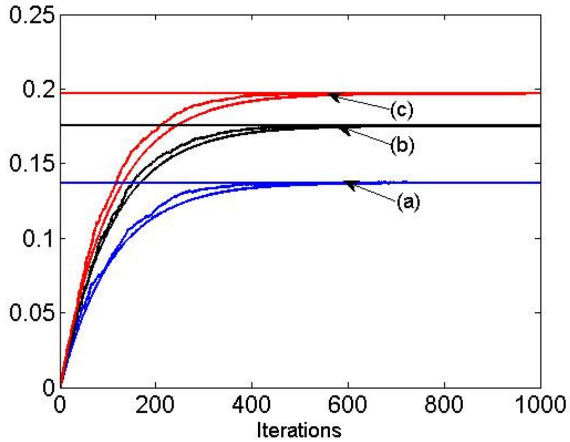

Figure 4 – Mean weight behavior for P=4 and α=0.4. Simulations (ragged) and model (continuous). (a) coefficient 20; (b) coefficient 40; (c) coefficient 30. Horizontal lines represent the true values of the plant coefficients.

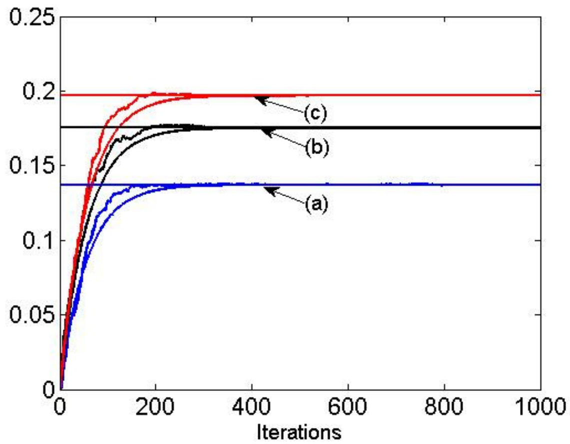

Figure 5 – Mean weight behavior for P=2 and α=0.9. Simulations (ragged) and model (continuous). (a) coefficient 20; (b) coefficient 40; (c) coefficient 30. Horizontal lines represent the true values of the plant coefficients.

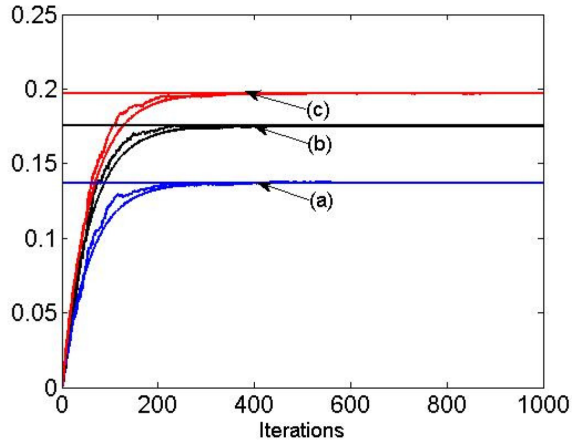

Figure 6 – Mean weight behavior for P=4 and α=0.9. Simulations (ragged) and model (continuous). (a) coefficient 20; (b) coefficient 40; (c) coefficient 30. Horizontal lines represent the true values of the plant coefficients.