16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

AN ADAPTIVE ALGORITHM FOR HAMMERSTEIN FILTER SYSTEM IDENTIFICATION

Ifiok J. Umoh and Tokunbo Ogunfunmi

Department of Electrical Engineering, Santa Clara University
500 ElCamino Real, Santa Clara, CA 95053, U.S.A

ABSTRACT

This paper develops an affine projection-based variable step-size adaptive algorithm for updating the coefficients of a Hammerstein nonlinear system. The variable step-size adaptive Hammerstein algorithm, recursively updates the step-size using a new criterion that provides a measure of how close the Hammerstein filter is to optimum performance. Experimental results show that the algorithm converges to the true unknown system in the mean square sense. Results also show the superior convergence of our proposed algorithm in comparison to existing adaptive Hammerstein algorithms.

1. INTRODUCTION

The Hammerstein system model is a kind of nonlinear system comprising a cascade connection of a memoryless nonlinearity in series with a linear system as shown in Figure 1. This kind of nonlinear system has been used in modeling distillation columns [1], electrical drives[2], power amplifiers[3], isometric muscles[4], etc. Other nonlinear models such as the Wiener and Volterra models also exist and are used in modeling various nonlinear systems[5].

Some existing methods for Hammerstein system identification are the iterative[6], over-parameterization[7] stochastic[8,10], nonlinear least squares[11], separable least squares [12] and blind[13] methods. For the over-parameterization method, adaptive algorithms have been developed [7]. The resulting over-parameterized system from this method is linear and thus any linear adaptive algorithm can be applied in the estimation of the nonlinear Hammerstein system coefficients. The limitation of the over-parameterization approach is that the dimension of the resulting linear system is large, leading to convergence and robustness problems in the presence of noise. In [10] the stochastic gradient method is used in developing an adaptive algorithm to adjust the parameters of a pre-compensator. An online identification method for the Hammerstein model, based on the Kalman filter was designed in [14]. The authors assumed the parameters of the nonlinear system to be constant. This assumption was shown in [9] to yield the problem of slow convergence and high misadjustment in the presence of correlated inputs. In [9] the authors also proposed a stochastic gradient based adaptive Hammerstein filter algorithm which out-performs [14]. The stochastic method proposed in [8] has the advantage of not requiring prior knowledge of the form of the unknown nonlinearity. It's application is limited since most stochastic methods require the input to be white.

In this paper, we use the idea of stochastic gradient in [9], [15] and [16] to propose an affine projection based variable step size adaptive algorithm to estimate the linear and nonlinear coefficients of the Hammerstein system. Also, a new criterion that provides an indication of how close the system is to optimal performance is presented and used to recursively update the step size. This recursion is in such a way as to guarantee the stability of the Hammerstein system. The proposed algorithm provides a faster convergence rate compared to existing adaptive Hammerstein algorithms.

This work was supported by the Edward M Fellows Foundation scholarship at Santa Clara University

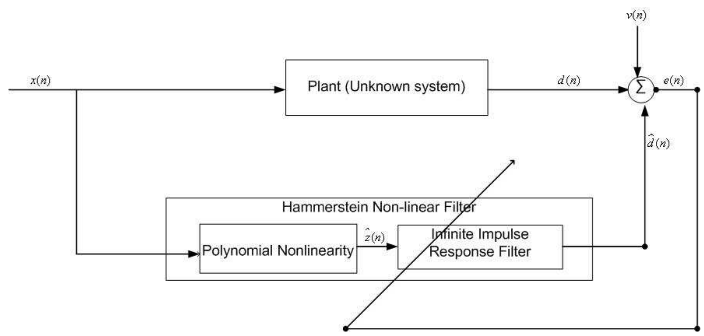

Figure 1: Adaptive system identification of a Hammerstein system model

2. PROBLEM STATEMENT

Consider the Hammerstein model shown in Figure 1, where x(n), v(n) and $\hat{d}(n)$ are the systems input, noise and output respectively. $\hat{z}(n)$ represents the unavailable internal signal output of the memoryless polynomial nonlinear system. The output of the memoryless polynomial nonlinear system, which is the input to the linear system is given by

$$\hat{z}(n) = \sum_{l=1}^{L} p_l(n)x^l(n) \tag{1}$$

Let the discrete linear time-invariant system be an infinite-impulse response (IIR) filter satisfying a linear difference equation of the form

$$\hat{d}(n) = -\sum_{i=1}^{N} a_i(n)\hat{d}(n-i) + \sum_{j=0}^{M} b_j(n)\hat{z}(n-j) \tag{2}$$

where $p_l(n)$, $a_i(n)$ and $b_j(n)$ represent the coefficients of the nonlinear Hammerstein system at any given time n. To ensure uniqueness, we set $b_0(n) = 1$ (any other coefficient other than $b_0(n)$ can be set to 1). Thus, (2) can be written as

$$\hat{d}(n) = \sum_{l=1}^{L} p_l(n) x^l(n) - \sum_{i=1}^{N} a_i(n)\hat{d}(n-i) + \sum_{j=1}^{M} b_j(n)\hat{z}(n-j) \tag{3}$$

Let

$$\hat{\theta}(n) = [\ a_1(n) \quad \dots \quad a_N(n) \quad b_1(n) \quad \dots \quad b_M(n)$$
$$p_1(n) \quad \dots \quad p_L(n)\ ]^H$$

$$b_0 = 1$$

$$\hat{\mathrm{s}}(n) = [\ -\hat{d}(n-1) \quad \dots \quad -\hat{d}(n-N) \quad \hat{z}(n-1) \quad \dots \quad \hat{z}(n-M)$$
$$x(n) \quad \dots \quad x^L(n)\ ]^H$$

Equation (3) can be rewritten in compact form

$$\hat{d}(n) = \hat{\mathrm{s}}(n)^H \hat{\theta}(n) \tag{4}$$

The goal of the Adaptive Hammerstein system identification is to update the coefficient vector ($\hat{\theta}(n)$) in (4) of the nonlinear Hammerstein filter based only on the input signal $x(n)$ and output signal $d(n)$ such that $\hat{d}(n)$ is close to the desired response signal $d(n)$.

3. ADAPTIVE HAMMERSTEIN ALGORITHM

In this section, we develop an algorithm based on the theory of Affine projection for estimation of the coefficients of the nonlinear Hammerstein system using the plant input and output signals. The main idea of our approach to nonlinear Hammerstein system identification is to formulate a criterion for designing a variable step-size affine projection Hammerstein filter algorithm and then use the criterion in minimizing the cost function.

3.1 Stochastic Gradient Minimization approach

We formulate the criterion for designing the adaptive Hammerstein filter as the minimization of the square Euclidean norm of the change in the weight vector

$$\tilde{\theta}(n) = \hat{\theta}(n) - \hat{\theta}(n-1) \tag{5}$$

subject to the set of Q constraints

$$d(n-q) = \hat{\mathrm{s}}(n-q)^H \hat{\theta}(n) \quad q = 1, \dots, \mathrm{Q} \tag{6}$$

Applying the method of Lagrange multipliers with multiple constraints to (5) and (6), the cost function for the affine projection filter is written as

$$J(n-1) = \left\| \hat{\theta}(n) - \hat{\theta}(n-1) \right\|^2 + \mathrm{Re}\left[\varepsilon(n-1)\lambda\right] \tag{7}$$

where

$$\varepsilon(n-1) = \mathbf{d}(n-1) - \hat{S}(n-1)^H \hat{\theta}(n) \tag{8}$$

$$\mathbf{d}(n-1) = [\ d(n-1) \quad \dots \quad d(n-Q)\ ]^H$$

$$\hat{S}(n-1) = [\ \hat{\mathrm{s}}(n-1) \quad \dots \quad \hat{\mathrm{s}}(n-Q)\ ]$$

$$\lambda = [\ \lambda_1 \quad \dots \quad \lambda_Q\ ]^H$$

Minimizing the cost function (7) (squared prediction error) with respect to the nonlinear Hammerstein filter weight vector $\hat{\theta}(n)$ gives

$$\frac{\partial J(n-1)}{\partial \hat{\theta}(n)} = 2\left(\hat{\theta}(n) - \hat{\theta}(n-1)\right) - \frac{\partial\left(\hat{\theta}(n)^H \hat{S}(n-1)\right)\lambda}{\partial \hat{\theta}(n)} \tag{9}$$

where

$$\frac{\partial\left(\hat{\theta}(n)^H \hat{S}(n-1)\right)}{\partial \hat{\theta}(n)} = \left[\ \frac{\partial \hat{\theta}(n)^H \hat{\mathrm{s}}(n-1)}{\partial \hat{\theta}(n)} \quad \dots \quad \frac{\partial \hat{\theta}(n)^H \hat{\mathrm{s}}(n-Q)}{\partial \hat{\theta}(n)}\ \right] \tag{10}$$

Since a portion of the vectors $\hat{s}(n)$ in $\hat{S}(n)$ include past $\hat{d}(n)$ which are dependent on past $\hat{\theta}(n)$ which are used to form the new $\hat{\theta}(n)$, the partial derivative of each element in (9) gives

$$\frac{\partial \hat{\theta}(n)^H \hat{\mathrm{s}}(n-q)}{\partial a_i(n)} = -\hat{d}(n-q-i) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial a_i(n)} \tag{11}$$

$$\frac{\partial \hat{\theta}(n)^H \hat{\mathrm{s}}(n-q)}{\partial b_j(n)} = \hat{z}(n-q-j) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial b_j(n)} \tag{12}$$

$$\frac{\partial \hat{\theta}(n)^H \hat{\mathrm{s}}(n-q)}{\partial p_l(n)} = x^l(n-q) + \sum_{k=1}^{M} b_k(n) \frac{\partial \hat{z}(n-q-k)}{\partial p_l(n)} - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial p_l(n)} \tag{13}$$

From (11), (12) and (13) it is necessary to evaluate the derivative of past $\hat{d}(n)$ with respect to current weight estimates. In evaluating the derivative of $\hat{d}(n)$ with respect to the current weight vector, we assume that the step size of the adaptive algorithm is chosen such that [16]

$$\hat{\theta}(n) \cong \hat{\theta}(n-1) \cong \dots \cong \hat{\theta}(n-N)$$

Therefore

$$a_i(n) \cong a_i(n-1) \cong \dots \cong a_i(n-N)$$

$$\frac{\partial \hat{d}(n-q)}{\partial a_i(n)} = -\hat{d}(n-q-i) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial a_i(n-k)} \tag{14}$$

$$b_j(n) \cong b_j(n-1) \cong \dots \cong b_j(n-N)$$

$$\frac{\partial \hat{d}(n-q)}{\partial b_j(n)} = \hat{z}(n-q-j) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial b_j(n-k)} \tag{15}$$

$$p_l(n) \cong p_l(n-1) \cong \dots \cong p_l(n-N)$$

and
$$\frac{\partial p_l(n-q-k)}{\partial p_l(n-k)} = 1$$
thus,
$$\frac{\partial \hat{d}(n-q)}{\partial p_l(n)} = x^l(n-q) + \sum_{k=1}^{M} b_k(n) x^l(n-q-k) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial p_l(n-k)} \quad (16)$$
where
$$\hat{\phi}(n-q) = \frac{\partial \hat{d}(n-q)}{\partial \hat{\theta}(n)} = \left[\begin{array}{cccccccc} \frac{\partial \hat{d}(n-q)}{\partial a_1(n)} & \ldots & \frac{\partial \hat{d}(n-q)}{\partial a_N(n)} & \frac{\partial \hat{d}(n-q)}{\partial b_1(n)} & \ldots & \frac{\partial \hat{d}(n-q)}{\partial b_M(n)} & \frac{\partial \hat{d}(n-q)}{\partial p_1(n)} & \ldots & \frac{\partial \hat{d}(n-q)}{\partial p_L(n)} \end{array} \right]^H$$
Let
$$\hat{\Phi}(n-1) = \frac{\partial \left(\hat{\theta}(n)^H \hat{S}(n-1) \right)}{\partial \hat{\theta}(n)}$$

$$\hat{\psi}(n-q) = \left[\begin{array}{c} -\hat{d}(n-q-1) \quad \ldots \quad -\hat{d}(n-q-N) \\ \hat{z}(n-q-1) \quad \ldots \quad \hat{z}(n-q-M) \\ \sum_{j=0}^{M} x(n-q-j) \quad \ldots \quad \sum_{j=0}^{M} x^L(n-q-j) \end{array} \right]^H$$

$$\hat{\Psi}(n-1) = \left[\begin{array}{ccc} \hat{\psi}(n-1) & \ldots & \hat{\psi}(n-Q) \end{array} \right]$$
Substituting (14),(15) and (16) into (10), we get
$$\hat{\Phi}(n-1) = \hat{\Psi}(n-1) - \sum_{k=1}^{N} a_k(n-1) \hat{\Phi}(n-1-k) \quad (17)$$
Thus, rewriting (9)
$$\frac{\partial J(n-1)}{\partial \hat{\theta}(n)} = 2\left(\hat{\theta}(n) - \hat{\theta}(n-1)\right) - \hat{\Phi}(n-1)\lambda \quad (18)$$
Setting the partial derivative of the cost function in (18) to zero, we get
$$\tilde{\theta}(n) = \frac{1}{2} \hat{\Phi}(n-1)\lambda \quad (19)$$
From (6), we can write
$$\mathbf{d}(n-1) = \hat{S}(n-1)^H \hat{\theta}(n) \quad (20)$$
where
$$\mathbf{d}(n-1) = \left[\begin{array}{ccc} d(n-1) & \ldots & d(n-Q) \end{array} \right]$$

$$\mathbf{d}(n-1) = \hat{S}(n-1)^H \hat{\theta}(n-1) + \frac{1}{2} \hat{S}(n-1)^H \hat{\Phi}(n-1)\lambda \quad (21)$$
Evaluating (21) for λ results in
$$\lambda = 2\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1) \quad (22)$$
where
$$\mathbf{e}(n-1) = \mathbf{d}(n-1) - \hat{S}(n-1)^H \hat{\theta}(n-1)$$
Substituting (22) into (19) yields the optimum change in the weight vector
$$\tilde{\theta}(n) = \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1) \quad (23)$$

Assuming that the input to the linear part of the nonlinear Hammerstein filter is a memoryless polynomial nonlinearity, we normalize (23) as in [17] and exercise control over the change in the weight vector from one iteration to the next keeping the same direction by introducing the step size μ. Regularization of the $\hat{S}(n-1)^H \hat{\Phi}(n-1)$ matrix is also used to guard against numerical difficulties during inversion. Thus yielding

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu \hat{\Phi}(n-1)\left(\delta I + \mu \hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1) \quad (24)$$

To improve the update process Newton's method is applied by scaling the update vector by $R^{-1}(n)$. The matrix $R(n)$ is recursively computed as
$$R(n) = \lambda_n R(n-1) + (1-\lambda_n)\hat{\Phi}(n-1)\hat{\Phi}(n-1)^H \quad (25)$$
where λ_n is typically chosen between 0.95 and 0.99. Applying the matrix inversion lemma on (25) and using the result in (24), the new update equation is given by

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu R(n-1)^{-1} \hat{\Phi}(n-1)\left(\delta I + \mu \hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1) \quad (26)$$

3.2 Variable Step-Size

In this subsection, we derive a variable step-size by using a Lyapunov function for the Hammerstein filter weight update. The variable step size derived guarantees the stable operation of the linear filter by satisfying the condition for the choice of μ in [18]. Let
$$\bar{\theta}(n) = \theta - \hat{\theta}(n)$$
We propose the Lyapunov function $V(n)$ as
$$V(n) = \bar{\theta}(n)^H \bar{\theta}(n) \quad (27)$$
which is the general form of the quadratic Lyapunov function. The Lyapunov function is positive definite in a range of values close to the optimum $\theta = \hat{\theta}(n)$. In order for the multidimensional error surface to be concave, the time derivative of the Lyapunov function must be semidefinite. This implies that
$$\Delta V(n) = V(n) - V(n-1) \leq 0 \quad (28)$$
From the Hammerstein filter update equation
$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1) \quad (29)$$
we subtract θ from both sides to yeild
$$\bar{\theta}(n) = \bar{\theta}(n-1) - \mu \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1) \quad (30)$$

From (27),(28) and (30) we have

$$\Delta V(n) = \left(\bar{\theta}(n)\right)^H \bar{\theta}(n) - \left(\bar{\theta}(n-1)\right)^H \bar{\theta}(n-1)$$

Minimizing the Lyapunov fuction with respect to the step-size μ and equating the result to zero, we obtain the optimum value for μ as μ_{opt}

$$\mu_{opt} = \frac{E\left[\bar{\theta}(n-1)^H \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1)\right]}{E\left[\mathbf{e}(n-1)^H \Upsilon(n-1)^H \Upsilon(n-1)\mathbf{e}(n-1)\right]} \quad (31)$$

were

$$\Upsilon(n-1) = \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \quad (32)$$

Adding the system noise $\mathbf{v}(n)$ to the desired output and assuming that the noise is independently and identically distributed and statistically independent of $\hat{S}(n)$. We have

$$\mathbf{d}(n) = \hat{S}(n)^H \theta + \mathbf{v}(n) \quad (33)$$

From (31) we write

$$\mu_{opt} E\left[\mathbf{e}(n-1)^H \Upsilon(n-1)^H \Upsilon(n-1)\mathbf{e}(n-1)\right] = E\left[\bar{\theta}(n-1)^H \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1)\right] \quad (34)$$

Thus, we propose the following suboptimal estimate for μ

$$\mu(n) = \frac{\hat{\mu}_{opt} E\left\|\Upsilon(n-1)\mathbf{e}(n-1)\right\|^2}{E\left\|\Upsilon(n-1)\mathbf{e}(n-1)\right\|^2 + \sigma_v^2 Tr\left\{E\left\|\Upsilon(n-1)\right\|^2\right\}}$$

We estimate $E\left\|\Upsilon(n-1)\mathbf{e}(n-1)\right\|$ by time averaging as follows

$$\hat{B}(n) = \alpha \hat{B}(n-1) - (1-\alpha)\Upsilon(n-1)\mathbf{e}(n-1)$$

$$\mu(n) = \hat{\mu}_{opt}\left(\frac{\left\|\hat{B}(n)\right\|^2}{\left\|\hat{B}(n)\right\|^2 + C}\right) \quad (35)$$

where $\hat{\mu}_{opt}$ is an rough estimate of μ_{opt} and C is a constant representing $\sigma_v^2 Tr\left\{\left\|\Upsilon(n-1)\right\|^2\right\} \approx \frac{Q}{SNR}$. We guarantee the stability of the Hammerstein filter by choosing $\hat{\mu}_{opt}$ to satisfy [18].

Algorithm 1 Summary of the proposed Variable Stepsize Hammerstein adaptive algorithm

INITIALIZE: $R^{-1}(0) = I, \lambda_n \neq 0, 0 < \beta \leq 1$
for $n = 0$ to sample size **do**

$\mathbf{e}(n-1) = \mathbf{d}(n-1) - \hat{S}(n-1)^H \hat{\theta}(n-1)$

$\hat{\Phi}(n-1) = \hat{\Psi}(n-1) - \sum_{k=1}^{N} a_k(n-1)\hat{\Phi}(n-1-k)$

$\hat{B}(n) = \alpha \hat{B}(n-1) - (1-\alpha)\Upsilon(n-1)\mathbf{e}(n-1)$

$\mu(n) = \hat{\mu}_{opt}\left(\frac{\|\hat{B}(n)\|^2}{\|\hat{B}(n)\|^2 + C}\right)$

$R(n)^{-1} = \frac{1}{\lambda_n}\Big[R(n-1)^{-1} - R(n-1)^{-1}\hat{\Phi}(n-1) \left(\frac{\lambda_n}{1-\lambda_n} I - \hat{\Phi}(n-1)^H R(n-1)^{-1}\hat{\Phi}(n-1)\right)^{-1} \hat{\Phi}(n-1)^H R(n-1)^{-1}\Big]$

$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu R(n)^{-1}\hat{\Phi}(n-1) \left(\delta I + \mu \hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1)$

$\hat{z}(n) = \mathbf{x}(n)^H \hat{\mathbf{p}}(n)$

$\hat{d}(n) = \hat{\mathbf{s}}(n)^H \hat{\theta}(n)$

end for

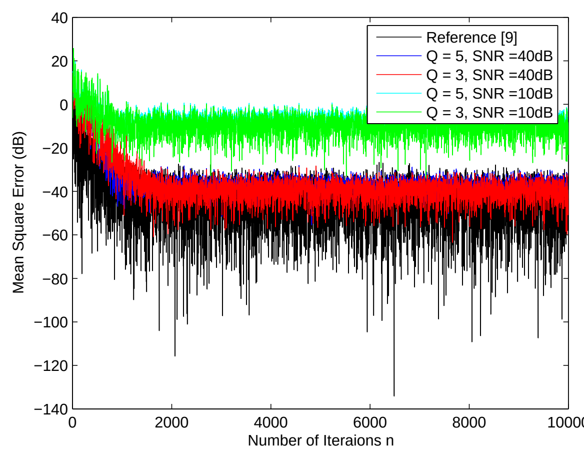

Figure 2: Mean square Error of the proposed algorithm given a white Gaussian input of variance 0.5

4. SIMULATION RESULTS

In this section, we illustrate the performance of the proposed nonlinear Hammerstein algorithm by applying the algorithm to the identification of an unknown Hammerstein system as shown in Figure 1. The unknown system was characterized to have a memoryless nonlinearity with input-output relationship given by

$$z(n) = 0.4x(n) - 0.3x(n)^2 + 0.2x(n)^3 \quad (36)$$

and the transfer function of the linear part was given by

$$H_2(z) = \frac{1 - 4z^{-1} + 6z^{-2} - 4z^{-3} + 1z^{-4}}{1 - 0.3906z^{-1} + 0.5343z^{-2} - 0.0842z^{-3} + 0.0207z^{-4}} \quad (37)$$

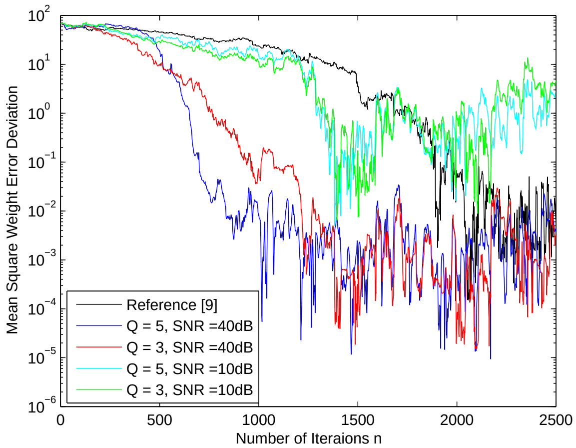

Figure 3: Mean square Weight deviation of the proposed algorithm given a white Gaussian input of variance 0.5

The unknown system was excited by a zero mean, white Gaussian signal with variance of 0.5. The desired response signal $d(n)$ of the adaptive filter was obtained by corrupting the output of the unknown system with additive white Gaussian noise with variance such that the output signal to noise ratio (SNR) was 10dB for low SNR and 40dB for high SNR experiments. The proposed adaptive filter was implemented with the parameter $\hat{\mu}_{opt}$ set to $1 \cdot 10^{-5}$ in the low SNR experiment and $2.5 \cdot 10^{-5}$ in the high SNR experiment, C set to 0.0001, $\delta = 10^{-3}$, $\lambda_n = 0.95$ and results were obtained by averaging over 100 independent runs.

Figure 2 shows that the mean square error performance for the proposed algorithm under both low and high SNR. The figure also shows a comparison between our proposed algorithm and [9] with SNR set to 40dB. The result shows that the mean square error performance of both algorithms are comparable. Considering the evolution of the mean square weight deviation from it's true value shown in Figure 3, our results indicate that at low SNR, the mean square weight deviation of the algorithm is not dependent on the block size Q. While at high SNR, the convergence rate increases with an increase in the block size Q. It also shows clear our proposed algorithm has a faster convergence rate compared to [9].

5. CONCLUSION

We have proposed a new adaptive filtering algorithm for the Hammerstein model filter based on the theory of Affine Projections. The new algorithm uses the norm of the projected weight error vector as a criterion to track the adaptive Hammerstein algorithm's optimum performance. Simulation result confirm the algorithms convergence to it's true unknown system in the mean square sense faster than existing algorithms.

REFERENCES

[1] R. K. Pearson and M. Pottmann, "Gray-box identification of block-oriented nonlinear models," *J. Process. Contr.*, vol. 10, no. 4, pp. 301–315, 2000.

[2] A. Balestrino, A. Landi, M. Ould-Zmirli and L. Sani, "Automatic nonlinear auto-tuning method for Hammerstein modeling of electical drives," *IEEE Trans. Ind. Electron.*, vol. 48, pp. 645–655, 2001.

[3] H. W. Kang, Y. S. Cho and D. H. Youn, "On compensating nonlinear distortions of an OFDM system using an efficient adaptive predistorter," *IEEE Trans. Commun.*, vol. 47, no. 4, pp. 522–526, 1999.

[4] K. J. Hunt, M. Munih, N. de. N. Donaldson and F. M. D. Barr, "Investigation of the Hammerstein hypothesis in the modeling of electrically stimulated muscle," *IEEE Trans. Biomed. Eng.*, vol. 45, no. 8, pp. 998–1009, 1998.

[5] T. Ogunfunmi, *Adaptive nonlinear system identification: Volterra and Wiener model approaches*. Springer-Verlag, September 2007.

[6] J. Voros, "Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities," *IEEE Trans. Autom. Control*, vol. 44, no. 11, pp. 2145–2149, November 1999.

[7] F. Ding, Y. Shi and T. Chen, "Gradient based identification methods for Hammerstein nonlinear ARMAX models," *Nonlinear Dynamics*, vol. 45, pp. 31–43, 2005.

[8] W. Greblicki, "Continuous-time Hammerstein system identification," *IEEE Trans. Autom. Control*, vol. 45, pp. 1232–1236, June 2000.

[9] J. Jeraj and V. J. Mathews, "A stable adaptive Hammerstein filter employing partial orthogonalization of the input signals," *IEEE Trans. Signal Process.*, vol. 54, pp. 1412–1420, April 2006.

[10] H. W. Kang, Y. S. Cho and D. H. Youn, "Adaptive precompensation of Wiener systems," *IEEE Trans. Signal Process.*, vol. 46, pp. 2825–2829, October 1998.

[11] E. W. Bai, "Identification of linear systems with hard input nonlinearities of known structure," *Automatica*, vol. 38, pp. 853–860, 2002.

[12] E. Bauer and B. Ninness, "Asymptotic properties of Hammerstein model estimates," in *Proc. IEEE Conf. Decision Control*, Sydney, Australia, 2000, pp. 2855–2860.

[13] E. W. Bai, "A blind approach to the Hammerstein-Wiener model identification," *Automatica*, vol. 38, pp. 967–976, 2002.

[14] M.Boutayeb and M. Darouach, "Recursive identification method for the Hammerstein model: Extension to nonlinear MISO systems," *Control Theory Adv. Technol.*, vol. 10, no. 1, pp. 57–72, 1994.

[15] H. Shin, A. H. Sayed and W. Song, "Variable step-size NLMS and affine projection algorithms," *IEEE Signal Process. Letters*, vol. 11, no. 2, pp. 132 – 135, Febrary 2004.

[16] S. Haykin, *Adaptive Filter Theory*, 4th ed. New Jersey, USA: Prentice Hall, 2002.

[17] N. J. Bershad, "On the optimum data nonlinearity of LMS adaptation," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 34, pp. 69 – 76, Feburary 1986.

[18] A. Carini, V. J. Mathews, and G. L. Sicuranza", "Sufficient stability bounds for slowly varying direct-form recursive linear filters and their application in adaptive iir filters," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2561 – 2567, September 1999.