Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2009, Article ID 859698, 7 pages
doi:10.1155/2009/859698

Research Article

An Adaptive Nonlinear Filter for System Identification

Ifiok J. Umoh (EURASIP Member) and Tokunbo Ogunfunmi

Department of Electrical Engineering, Santa Clara University, Santa Clara, CA 95053, USA

Correspondence should be addressed to Tokunbo Ogunfunmi, togunfunmi@scu.edu

Received 12 March 2009; Accepted 8 May 2009

Recommended by Jonathon Chambers

The primary difficulty in the identification of Hammerstein nonlinear systems (a static memoryless nonlinear system in series with a dynamic linear system) is that the output of the nonlinear system (input to the linear system) is unknown. By employing the theory of affine projection, we propose a gradient-based adaptive Hammerstein algorithm with variable step-size which estimates the Hammerstein nonlinear system parameters. The adaptive Hammerstein nonlinear system parameter estimation algorithm proposed is accomplished without linearizing the systems nonlinearity. To reduce the effects of eigenvalue spread as a result of the Hammerstein system nonlinearity, a new criterion that provides a measure of how close the Hammerstein filter is to optimum performance was used to update the step-size. Experimental results are presented to validate our proposed variable step-size adaptive Hammerstein algorithm given a real life system and a hypothetical case.

Copyright © 2009 I. J. Umoh and T. Ogunfunmi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Nonlinear system identification has been an area of active research for decades. Nonlinear systems research has led to the discovery of numerous types of nonlinear systems such as Volterra, Hammerstein, and Weiner nonlinear systems [1–4]. This work will focus on the Hammerstein nonlinear system depicted in Figure 1. Hammerstein nonlinear models have been applied to modeling distortion in nonlinearly amplified digital communication signals (satellite and microwave links) followed by a linear channel [5, 6]. In the area of biomedical engineering, the Hammerstein model finds application in modeling the involuntary contraction of human muscles [7, 8] and human heart rate regulation during treadmill exercise [9]. Hammerstein systems are also applied in the area of Neural Network since it provides a convenient way to deal with nonlinearity [10]. Existing Hammerstein nonlinear system identification techniques can be divided into three groups:

(i) deterministic techniques such as orthogonal least-squares expansion method [11–13],

(ii) stochastic techniques based on recursive algorithms [14, 15] or nonadaptive methods [16], and

(iii) adaptive techniques [17–20].

Adaptive Hammerstein algorithms have been achieved using block based adaptive algorithms [11, 20]. In block based adaptive Hammerstein algorithms, the Hammerstein system is overparameterized in such a way that the Hammerstein system is linear in the unknown parameters. This allows the use of any linear estimation algorithm in solving the Hammerstein nonlinear system identification problem. The limitation of this approach is that the dimension of the resulting linear block system can be very large, and therefore, convergence or robustness of the algorithm becomes an issue [18]. Recently, Bai reported a blind approach to Hammerstein system identification using least mean square (LMS) algorithm [18]. The method reported applied a two-stage identification process (Linear Infinite Impulse Response (IIR) stage and the nonlinear stage) without any knowledge of the internal signals connecting both cascades in the Hammerstein system. This method requires a white input signal to guarantee the stability and convergence of the algorithm. Jeraj and Mathews derived an adaptive Hammerstein system identification algorithm by linearizing the system nonlinearity using a Gram-Schmidt orthogonalizer at the input to the linear subsystem (forming an MISO system) [17]. This method also suffers the same limitations as the block-based adaptive Hammerstein algorithms. Thus, to improve the speed of

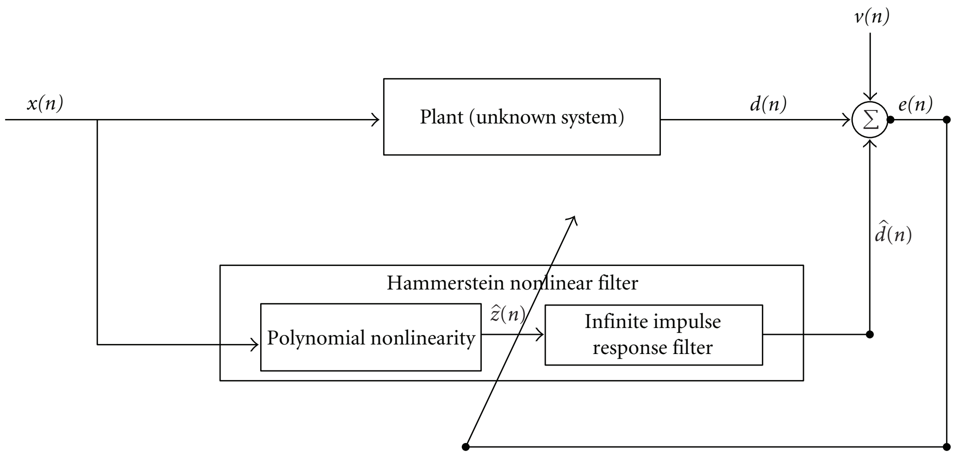

Figure 1: Adaptive system identification of a Hammerstein system model.

convergence while maintaining a small misadjustment and computational complexity, the Affine Projection theory is used as opposed to LMS [18] or Recursive Least squares (RLSs).

In nonlinear system identification, input signals with high eigen value spread, ill-conditioned tap input autocorrelation matrix can lead to divergence or poor performance of a fixed step-size adaptive algorithm. To mitigate this problem, a number of variable step-size update algorithms have been proposed. These variable step-size update algorithms can be roughly divided into gradient adaptive step-size [21, 22] and normalized generalized gradient descent [23]. The major limitation of gradient adaptive step-size algorithms is their sensitivity to the time correlation between input signal samples and the value of the additional step-size parameter that governs the gradient adaptation of the step-size. As a result of these limitations, a criteria for the choice of the step-size based on Lyapunov stability theory is proposed to track the optimal step-size required to maintain a fast convergence rate and low misadjustment.

In this paper, we focus on the adaptive system identification problem of a class of Hammerstein output error type nonlinear systems with polynomial nonlinearity. Our unique contributions in the paper are as follows.

(1) Using the theory of affine projections [24], we derive an adaptive Hammerstein algorithm that identifies the linear subsystem of the Hammerstein system without prior knowledge of the input signal $\hat{z}(n)$.

(2) Employing the Lyapunov stability theory, we develop criteria for the choice of the algorithms step-size which ensures the minimization of the Lyapunov function. This is particularly important for the stability of the linear algorithm regardless of the location of the poles of the IIR filter.

Briefly, the paper is organized as follows. Section 2 describes the nonlinear Hammerstein system identification problem addressed in this paper. Section 3 contains a detailed derivation of the proposed variable step-size adaptive Hammerstein algorithm. Section 4 provides both a hypothetical and real life data simulation validating the effectiveness of the variable step-size adaptive algorithm proposed. Finally, we conclude with a brief summary in Section 5.

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

$$\begin{aligned} \hat{\theta}(n) &= [a_1(n) \ \cdots \ a_N(n)\ b_1(n) \ \cdots \ b_M(n) \\ &\quad p_1(n) \ \cdots \ p_L(n)]^H, \\ b_0 &= 1, \\ \hat{s}(n) &= \Big[-\hat{d}(n-1) \ \cdots \ -\hat{d}(n-N) \\ &\quad \hat{z}(n-1) \ \cdots \ \hat{z}(n-M) \\ &\quad x(n) \ \cdots \ x^L(n)\Big]^H. \end{aligned} \tag{4}$$

Equation (3) can be rewritten in compact form

$$\hat{d}(n) = \hat{s}(n)^H \hat{\theta}(n). \tag{5}$$

The goal of the Adaptive nonlinear Hammerstein system identification is to estimate the coefficient vector ($\hat{\theta}(n)$) in (5) 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 [24] 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

$$\tilde{\theta}(n) = \hat{\theta}(n) - \hat{\theta}(n-1) \tag{6}$$

subject to the set of Q constraints

$$d(n-q) = \hat{s}(n-q)^H \hat{\theta}(n) \quad q = 1, \dots, Q. \tag{7}$$

Applying the method of Lagrange multipliers with multiple constraints to (6) and (7), the cost function for the affine projection filter is written as (assuming real data)

$$J(n-1) = \left\| \hat{\theta}(n) - \hat{\theta}(n-1) \right\|^2 + \operatorname{Re}[\epsilon(n-1)\lambda], \tag{8}$$

where

$$\begin{aligned} \epsilon(n-1) &= \mathbf{d}(n-1) - \hat{S}(n-1)^H \hat{\theta}(n), \\ \mathbf{d}(n-1) &= [d(n-1) \ \cdots \ d(n-Q)]^{\mathrm{H}}, \\ \hat{S}(n-1) &= [\hat{s}(n-1) \ \cdots \ \hat{s}(n-Q)], \\ \lambda &= [\lambda_1 \ \cdots \ \lambda_Q]^H. \end{aligned} \tag{9}$$

Minimizing the cost function (8) (squared prediction error) with respect to the nonlinear Hammerstein filter weight vector $\hat{\theta}(n)$ gives

$$\frac{\partial J(n-1)}{\partial \hat{\theta}(n)} = 2\left(\hat{\theta}(n) - \hat{\theta}(n-1)\right) - \frac{\partial\left(\hat{\theta}(n)^H \hat{S}(n-1)\right)\lambda}{\partial \hat{\theta}(n)}, \tag{10}$$

where

$$\begin{aligned} &\frac{\partial\left(\hat{\theta}(n)^H \hat{S}(n-1)\right)}{\partial \hat{\theta}(n)} \\ &\quad = \left[\frac{\partial \hat{\theta}(n)^H \hat{s}(n-1)}{\partial \hat{\theta}(n)} \ \cdots \ \frac{\partial \hat{\theta}(n)^H \hat{s}(n-Q)}{\partial \hat{\theta}(n)} \right]. \end{aligned} \tag{11}$$

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 (10) gives

$$\frac{\partial \hat{\theta}(n)^H \hat{s}(n-q)}{\partial a_i(n)} = -\hat{d}(n-q-i) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial a_i(n)}, \tag{12}$$

$$\frac{\partial \hat{\theta}(n)^H \hat{s}(n-q)}{\partial b_j(n)} = \hat{z}(n-q-j) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial b_j(n)}, \tag{13}$$

$$\begin{aligned} \frac{\partial \hat{\theta}(n)^H \hat{s}(n-q)}{\partial p_l(n)} &= x^l(n-q) + \sum_{k=1}^{M} b_k(n) \frac{\partial \hat{z}(n-q-k)}{\partial p_l(n)} \\ &\quad - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial p_l(n)}. \end{aligned} \tag{14}$$

From (12), (13), and (14) 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 [24]

$$\hat{\theta}(n) \cong \hat{\theta}(n-1) \cong \cdots \cong \hat{\theta}(n-N). \tag{15}$$

Therefore

$$\begin{aligned} &a_i(n) \cong a_i(n-1) \cong \cdots \cong a_i(n-N), \\ &\frac{\partial \hat{d}(n-q)}{\partial a_i(n)} = -\hat{d}(n-q-i) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial a_i(n-k)}, \\ &b_j(n) \cong b_j(n-1) \cong \cdots \cong b_j(n-N), \end{aligned} \tag{16}$$

$$\frac{\partial \hat{d}(n-q)}{\partial b_j(n)} = \hat{z}(n-q-j) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial b_j(n-k)},$$
$$p_l(n) \cong p_l(n-1) \cong \cdots \cong p_l(n-N), \tag{17}$$

$$\frac{\partial \hat{d}(n-q)}{\partial p_l(n)} = x^l(n-q) + \sum_{k=1}^{M} b_k(n) \frac{\partial \hat{z}(n-q-k)}{\partial p_l(n-k)}. - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial p_l(n-k)}, \tag{18}$$

$$\frac{\partial p_l(n-q-k)}{\partial p_l(n-k)} = 1, \tag{19}$$

thus,

$$\frac{\partial \hat{d}(n-q)}{\partial p_l(n)} = x^l(n-q) + \sum_{k=1}^{M} b_k(n) x^l(n-q-k) - \sum_{k=1}^{N} a_k(n) \frac{\partial \hat{d}(n-q-k)}{\partial p_l(n-k)}, \tag{20}$$

where

$$\hat{\phi}(n-q) = \frac{\partial \hat{d}(n-q)}{\partial \hat{\theta}(n)} = \left[\frac{\partial \hat{d}(n-q)}{\partial a_1(n)} \cdots \frac{\partial \hat{d}(n-q)}{\partial a_N(n)} \; \frac{\partial \hat{d}(n-q)}{\partial b_1(n)} \cdots \frac{\partial \hat{d}(n-q)}{\partial b_M(n)} \; \frac{\partial \hat{d}(n-q)}{\partial p_1(n)} \cdots \frac{\partial \hat{d}(n-q)}{\partial p_L(n)} \right]^H. \tag{21}$$

Let

$$\hat{\Phi}(n-1) = \frac{\partial\left(\hat{\theta}(n)^H \hat{S}(n-1)\right)}{\partial \hat{\theta}(n)},$$
$$\hat{\psi}(n-q) = \left[-\hat{d}(n-q-1) \cdots -\hat{d}(n-q-N) \; \hat{z}(n-q-1) \cdots \hat{z}(n-q-M) \; \sum_{j=0}^{M} x(n-q-j) \cdots \sum_{j=0}^{M} x^L(n-q-j) \right]^H,$$
$$\hat{\Psi}(n-1) = [\hat{\psi}(n-1) \cdots \hat{\psi}(n-Q)]. \tag{22}$$

Substituting (16), (17), and (20) into (11), we get

$$\hat{\Phi}(n-1) = \hat{\Psi}(n-1) - \sum_{k=1}^{N} a_k(n-1) \hat{\Phi}(n-1-k). \tag{23}$$

Thus, rewriting (10)

$$\frac{\partial J(n-1)}{\partial \hat{\theta}(n)} = 2\left(\hat{\theta}(n) - \hat{\theta}(n-1)\right) - \hat{\Phi}(n-1)\lambda. \tag{24}$$

Setting the partial derivative of the cost function in (24) to zero, we get

$$\tilde{\theta}(n) = \frac{1}{2} \hat{\Phi}(n-1)\lambda. \tag{25}$$

From (7), we can write

$$\mathbf{d}(n-1) = \hat{S}(n-1)^H \hat{\theta}(n), \tag{26}$$

where

$$\mathbf{d}(n-1) = [d(n-1) \cdots d(n-Q)],$$
$$\mathbf{d}(n-1) = \hat{S}(n-1)^H \hat{\theta}(n-1) + \frac{1}{2} \hat{S}(n-1)^H \hat{\Phi}(n-1)\lambda. \tag{27}$$

Evaluating (27) for λ results in

$$\lambda = 2\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1), \tag{28}$$

where

$$\mathbf{e}(n-1) = \mathbf{d}(n-1) - \hat{S}(n-1)^H \hat{\theta}(n-1). \tag{29}$$

Substituting (28) into (25) yields the optimum change in the weight vector

$$\tilde{\theta}(n) = \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} \mathbf{e}(n-1). \tag{30}$$

Assuming that the input to the linear part of the nonlinear Hammerstein filter is a memoryless polynomial nonlinearity, we normalize (30) as in [25] 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

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu \hat{\Phi}(n-1) \times \left(\hat{\Phi}\delta I + \mu \hat{S}(n-1)^H (n-1)\right)^{-1} \mathbf{e}(n-1). \tag{31}$$

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

$$R(n) = \lambda_n R(n-1) + (1-\lambda_n) \hat{\Phi}(n-1) \hat{\Phi}(n-1)^H, \tag{32}$$

where λ_n is typically chosen between 0.95 and 0.99. Applying the matrix inversion lemma on (32) and using the result in (31), the new update equation is given by

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu R(n-1)^{-1} \hat{\Phi}(n-1) \times \left(\delta I + \mu \hat{S}(n-1)^H \hat{\Phi}(n-1)\right)^{-1} e(n-1) \tag{33}$$

3.2. Variable Step-Size. In this subsection, we derive an update for the step-size using a Lyapunov function of summed squared nonlinear Hammerstein filter weight estimate error. The variable step-size derived guarantees the stable operation of the linear IIR filter by satisfying the stability condition for the choice of μ in [26]. Let

$$\overline{\theta}(n) = \theta - \hat{\theta}(n), \quad (34)$$

where θ represents the optimum Hammerstein system coefficient vector. We propose the Lyapunov function $V(n)$ as

$$V(n) = \overline{\theta}(n)^H \overline{\theta}(n), \quad (35)$$

which is the general form of the quadratic Lyapunov function [27]. 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

$$\Delta V(n) = V(n) - V(n-1) \leq 0. \quad (36)$$

From the Hammerstein filter update equation

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu\hat{\Phi}(n-1)\left(\hat{S}(n-1)^H\hat{\Phi}(n-1)\right)^{-1}\mathbf{e}(n-1), \quad (37)$$

we subtract θ from both sides to yeild

$$\overline{\theta}(n) = \overline{\theta}(n-1) - \mu\hat{\Phi}(n-1)\left(\hat{S}(n-1)^H\hat{\Phi}(n-1)\right)^{-1}\mathbf{e}(n-1). \quad (38)$$

From (35), (36), and (38) we have

$$\Delta V(n) = \left(\overline{\theta}(n)\right)^H\overline{\theta}(n) - \left(\overline{\theta}(n-1)\right)^H\overline{\theta}(n-1). \quad (39)$$

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

$$\mu_{\text{opt}} = \frac{E\left[\overline{\theta}(n-1)^H\hat{\Phi}(n-1)\left(\hat{S}(n-1)^H\hat{\Phi}(n-1)\right)^{-1}\mathbf{e}(n-1)\right]}{E\left[\mathbf{e}(n-1)^H\Upsilon(n-1)^H\Upsilon(n-1)\mathbf{e}(n-1)\right]}, \quad (40)$$

where

$$\Upsilon(n-1) = \hat{\Phi}(n-1)\left(\hat{S}(n-1)^H\hat{\Phi}(n-1)\right)^{-1}. \quad (41)$$

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

$$\mathbf{d}(n) = \hat{S}(n)^H\theta + \mathbf{v}(n). \quad (42)$$

INITIALIZE: $R^{-1}(0) = I, \lambda_n \neq 0, 0 < \beta \leq 1$
for $n = 0$ to sample size do
$\mathbf{e}(n-1) = \mathbf{d}(n-1) - \hat{S}(n-1)^H\hat{\theta}(n-1)$
$\hat{\Phi}(n-1) = \hat{\Psi}(n-1) - \sum_{k=1}^{N} a_k(n-1)\hat{\Phi}(n-1-k)$
$\hat{B}(n) = \alpha\hat{B}(n-1) - (1-\alpha)\Upsilon(n-1)\mathbf{e}(n-1)$
$\mu(n) = \hat{\mu}_{\text{opt}}(\frac{\|\hat{B}(n)\|^2}{\|\hat{B}(n)\|^2 + C})$
$(\frac{\lambda_n}{1-\lambda_n}I - \hat{\Phi}(n-1)^H R(n-1)^{-1}\hat{\Phi}(n-1))^{-1}$
$R(n)^{-1} = \frac{1}{\lambda_n}[R(n-1)^{-1} - R(n-1)^{-1}\hat{\Phi}(n-1)$
$\hat{\Phi}(n-1)^H R(n-1)^{-1}]$
$\hat{\theta}(n) = \hat{\theta}(n-1) - \mu(n)R(n)^{-1}\hat{\Phi}(n-1)$
$(\delta I + \mu(n)\hat{S}(n-1)^H\hat{\Phi}(n-1))^{-1}\mathbf{e}(n-1)$
$\hat{z}(n) = \mathbf{x}(n)^H\hat{\mathbf{p}}(n)$
$\hat{d}(n) = \hat{s}(n)^H\hat{\theta}(n)$
end for

Algorithm 1: Summary of the proposed Variable Step-size Hammerstein adaptive algorithm.

From (40) we write

$$\mu_{\text{opt}}E\left[\mathbf{e}(n-1)^H\Upsilon(n-1)^H\Upsilon(n-1)\mathbf{e}(n-1)\right] = E\left[\overline{\theta}(n-1)^H\hat{\Phi}(n-1)\left(\hat{S}(n-1)^H\hat{\Phi}(n-1)\right)^{-1}\mathbf{e}(n-1)\right]. \quad (43)$$

The computation of μ_{opt} requires the knowledge of $\overline{\theta}(n-1)$ which is not available during adaptation. Thus, we propose the following suboptimal estimate for $\mu(n)$:

$$\mu(n) = \frac{\hat{\mu}_{\text{opt}}E\|\Upsilon(n-1)\mathbf{e}(n-1)\|^2}{E\|\Upsilon(n-1)\mathbf{e}(n-1)\|^2 + \sigma_v^2 Tr\left\{E\|\Upsilon(n-1)\|^2\right\}}. \quad (44)$$

We estimate $E\|\Upsilon(n-1)e(n-1)\|$ by time averaging as follows:

$$\hat{B}(n) = \alpha\hat{B}(n-1) - (1-\alpha)\Upsilon(n-1)\mathbf{e}(n-1)$$
$$\mu(n) = \hat{\mu}_{\text{opt}}\left(\frac{\left\|\hat{B}(n)\right\|^2}{\left\|\hat{B}(n)\right\|^2 + C}\right), \quad (45)$$

where $\hat{\mu}_{\text{opt}}$ is an rough estimate of μ_{opt}, α is a smoothing factor ($0 < \alpha < 1$), and C is a constant representing $\sigma_v^2 Tr\{\|\Upsilon(n-1)\|^2\} \approx Q/\text{SNR}$. We guarantee the stability of the Hammerstein filter by choosing $\hat{\mu}_{\text{opt}}$ to satisfy the stability bound in [26]. Choosing $\hat{\mu}_{\text{opt}}$ to satisfy the stability bound [26] will bound the step-size update $\mu(n)$ with an upper limit of $\hat{\mu}_{\text{opt}}$ thereby ensuring the slow variation and stability of the linear IIR filter.

A summary of the proposed algorithm is shown in Algorithm 1. In the algorithm, N represents the number

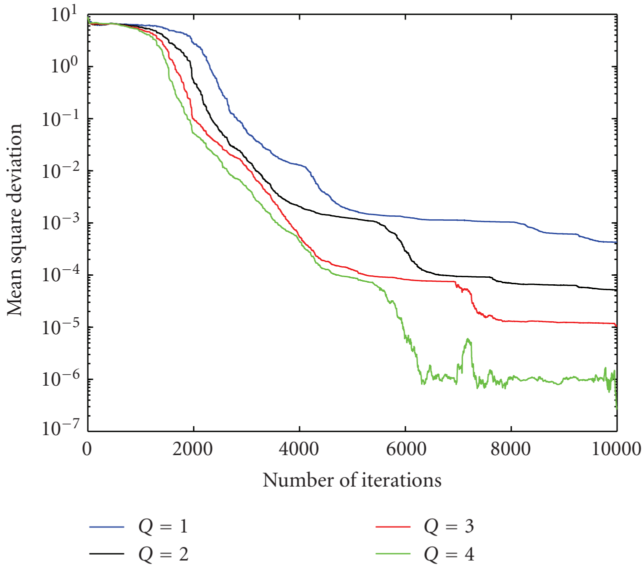

Figure 2: Mean square deviation learning curve of the proposed algorithm for varying constraint.

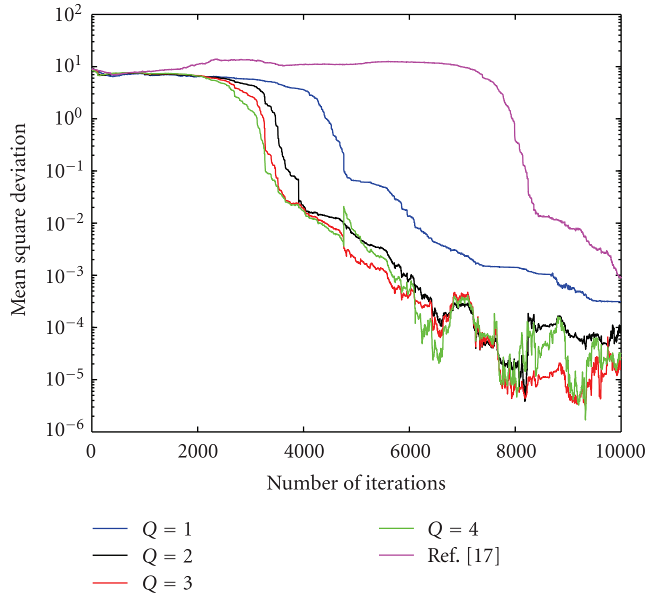

Figure 3: Mean square deviation learning curves of the proposed algorithm for varying constraint and a colored input signal.

of feedback, M the number of feedforward coefficients for the linear subsystem, and L the number of coefficients for the polynomial subsystem. Based on these coefficient numbers, let K represent $N + M + L - 2$ in the computation of the computational cost of our proposed adaptive nonlinear algorithm. In computing the computational cost, we assume that the cost of inverting a $K \times K$ matrix is $\bigcirc(K^3)$ (Multiplications and additions) and $\bigcirc(L^2N)$ for computing R^{-1} [28]. Under these assumptions, the computational cost of our proposed algorithm is of $\bigcirc(QK^2)$ multiplications compared to $\bigcirc(K^2)$ in [17]. This increase in complexity due to the order of the input regression matrix in the proposed algorithm is compensated for by the algorithms good performance.

4. Simulation Results

In this section, we validate the proposed algorithm with simulation results corresponding to two different types of input signals (white and highly colored signals). The white signal was an additive white Gaussian noise signal of zero mean and unity variance. The highly colored signal was generated by filtering the white signal with a filter of impulse response:

$$H_1(z) = \frac{0.5}{1 + 0.9z^{-1}}. \qquad (46)$$

The Hammerstein system was modeled such that the dynamic linear system had an impulse response $H_2(z)$ given by

$$H_2(z) = \frac{1.0000 - 1.8000z^{-1} + 1.6200z^{-2} - 1.4580z^{-3} + 0.6561z^{-4}}{1.0000 - 0.2314z^{-1} + 0.4318z^{-2} - 0.3404z^{-3} + 0.5184z^{-4}}, \qquad (47)$$

and static nonlinearity modeled as a polynomial with coefficients

$$z(n) = x(n) - 0.3x(n)^2 + 0.2x(n)^3. \qquad (48)$$

The desired response signal $d(n)$ of the adaptive Hammerstein filter was obtained by corrupting the output of the unknown system with additive white noise of zero mean and variance such that the output signal to noise ratio was 30 dB. The proposed algorithm was initialized as follows: $\lambda_n = 0.997$, $\mu_{\text{opt}} = 1.5e - 6$, $\delta = 5e - 4$, and $C = 0.0001$.

Figure 2 shows the mean square deviation between the estimated and optimum Hammerstein filter weights for a white input case. The results were obtained, by ensemble averaging over 100 independent trials. The figure shows that the convergence speed of the proposed algorithm is directly correlated to the number of constraints Q used in the algorithm.

Figure 3 shows a comparison of the mean square deviation learning curve obtained, from the proposed algorithm and the algorithm in [17]. The proposed algorithm outperformed [17] even though the authors had used the Gram Schmidt process to better enhance the algorithms performance in a colored environment. This result is expected since [17] is sensitive to the additional step-size and regularizing parameters that govern the adaptation of the Gram-Schmidt processor. Results also show that, above a constraint value of 1 ($Q > 1$), the proposed algorithms performance in the colored environment is similar.

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 minimizes the norm of the projected weight error vector as a criterion to track the adaptive Hammerstein algorithm's optimum performance. Simulation results confirm the convergence of the parameter estimates from our proposed algorithm to its corresponding parameter in the plants parameter vector.

References

[1] T. Ogunfunmi, *Adaptive Nonlinear System Identification: Volterra and Wiener Model Approaches*, Springer, London, UK, 2007.

[2] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*, John Wiley & Sons, New York, NY, USA, 2000.

[3] T. Ogunfunmi and S. L. Chang, "Second-order adaptive Volterra system identification based on discrete nonlinear Wiener model," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 148, no. 1, pp. 21–30, 2001.

[4] S.-L. Chang and T. Ogunfunmi, "Stochastic gradient based third-order Volterra system identification by using nonlinear Wiener adaptive algorithm," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 150, no. 2, pp. 90–98, 2003.

[5] W. Greblicki, "Nonlinearity estimation in Hammerstein systems based on ordered observations," *IEEE Transactions on Signal Processing*, vol. 44, no. 5, pp. 1224–1233, 1996.

[6] S. Prakriya and D. Hatzinakos, "Blind identification of linear subsystems of LTI-ZMNL-LTI models with cyclostationary inputs," *IEEE Transactions on Signal Processing*, vol. 45, no. 8, pp. 2023–2036, 1997.

[7] K. J. Hunt, M. Munih, N. N. de Donaldson, and F. M. D. Barr, "Investigation of the hammerstein hypothesis in the modeling of electrically stimulated muscle," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 8, pp. 998–1009, 1998.

[8] D. T. Westwick and R. E. Kearney, *Identification of Nonlinear Physiological Systems*, John Wiley & Sons, New York, NY, USA, 2003.

[9] S. W. Su, L. Wang, B. G. Celler, A. V. Savkin, and Y. Guo, "Identification and control for heart rate regulation during treadmill exercise," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 7, pp. 1238–1246, 2007.

[10] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, John Wiley & Sons, Chichester, UK, 2001.

[11] F. Ding, Y. Shi, and T. Chen, "Auxiliary model-based least-squares identification methods for Hammerstein output-error systems," *Systems and Control Letters*, vol. 56, no. 5, pp. 373–380, 2007.

[12] E.-W. Bai, "Identification of linear systems with hard input nonlinearities of known structure," *Automatica*, vol. 38, no. 5, pp. 853–860, 2002.

[13] F. Giri, F. Z. Chaoui, and Y. Rochdi, "Parameter identification of a class of Hammerstein plants," *Automatica*, vol. 37, no. 5, pp. 749–756, 2001.

[14] M. Boutayeb, H. Rafaralahy, and M. Darouach, "A robust and recursive indentificaition method for Hammerstein model," in *Proceedings of the IFAC World Congress*, pp. 447–452, San Francisco, Calif, USA, 1996.

[15] J. Vörös, "Iterative algorithm for parameter identification of Hammerstein systems with two-segment nonlinearities," *IEEE Transactions on Automatic Control*, vol. 44, no. 11, pp. 2145–2149, 1999.

[16] E.-W. Bai, "An optimal two-stage identification algorithm for Hammerstein-Wiener nonlinear systems," *Automatica*, vol. 34, no. 3, pp. 333–338, 1998.

[17] J. Jeraj and V. J. Mathews, "A stable adaptive Hammerstein filter employing partial orthogonalization of the input signals," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1412–1420, 2006.

[18] E.-W. Bai, "A blind approach to the Hammerstein-Wiener model identification," *Automatica*, vol. 38, no. 6, pp. 967–979, 2002.

[19] I. J. Umoh and T. Ogunfunmi, "An adaptive algorithm for Hammerstein filter system identification," in *Proceedings of the 16th European Signal Processing Conference (EUSIPCO '08)*, Lausanne, Switzerland, August 2008.

[20] F. Ding, T. Chen, and Z. Iwai, "Adaptive digital control of Hammerstein nonlinear systems with limited output sampling," *SIAM Journal on Control and Optimization*, vol. 45, no. 6, pp. 2257–2276, 2007.

[21] V. J. Mathews and Z. Xie, "Stochastic gradient adaptive filter with gradient adaptive step size," *IEEE Transactions on Signal Processing*, vol. 41, no. 6, pp. 2075–2087, 1993.

[22] A. Benveniste, M. Metivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximation*, Springer, New York, NY, USA, 1990.

[23] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Processing Letters*, vol. 11, no. 2, part 1, pp. 115–118, 2004.

[24] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 4th edition, 2002.

[25] N. J. Bershad, "On the optimum data nonlinearity in LMS adaptation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 1, pp. 69–76, 1986.

[26] A. Carini, V. John Mathews, and G. L. Sicuranza, "Sufficient stability bounds for slowly varying direct-form recursive linear filters and their applications in adaptive IIR filters," *IEEE Transactions on Signal Processing*, vol. 47, no. 9, pp. 2561–2567, 1999.

[27] C. R. Johnson Jr., "Adaptive IIR filtering: current results and open issues," *IEEE Transactions on Information Theory*, vol. 30, no. 2, pp. 237–250, 1984.

[28] J. M. Cioffi and T. Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 2, pp. 304–337, 1984.