

Parameter Estimation of Wiener Nonlinear Model Using Least Mean Square (LMS) Algorithm

Saurav Gupta^{*,1}, Ajit Kumar Sahoo^{*,2}, Upendra Kumar Sahoo^{*,3}

^{*}Department of Electronics and Communication Engineering

National Institute of Technology, Rourkela, India

Email: ¹greater.saurav@gmail.com, ²ajitsahoo@nitrkl.ac.in, ³sahooupen@nitrkl.ac.in

Abstract—In today's world of signal processing, nonlinear systems have been attained a considerable importance in the field of system identification and system control. The modeling of many physical systems was introduced by a nonlinear Wiener model consists of static nonlinear function followed by a linear time invariant (LTI) dynamic system. The output of the nonlinear function is considered to be continuous and invertible. This work leads the identification of Wiener model parameters using least mean square (LMS) algorithm and its two different variants named leaky LMS and modified leaky LMS due to its simple and effective adaptive nature. The simulation results for an example supporting the deduced methodology are obtained to effectively analyze the algorithm performance.

Index Terms—Signal processing, nonlinear, system identification, Wiener model, LMS.

I. INTRODUCTION

Parameter estimation is a fundamentally important task for controlling and identifying any system [1]. Nonlinear systems have received much attention over the last ten years in the field of fault detection [2], system control [3] and system identification [4]. For nonlinear systems, the block-oriented models are largely employed as their structures are quite simple and are also capable of approximating a wide range of nonlinear processes. These models consist of an interconnected linear time-invariant (LTI) system and a nonlinear static element [5]. Based on the position of nonlinear element, these models are classified as 1.) Wiener model, when a static nonlinear function is positioned following an LTI system and 2.) Hammerstein model, where a static nonlinear function is followed by an LTI system.

Wiener models are widely engaged in many engineering and science oriented practices. Some of the applications are modeling of pH neutralization process, electrical control, chemical processes and biological process. As far as the linear structured model is concerned, they can be easily derived from the block-structured nonlinear model. Hence, compared to general nonlinear model, controlling and optimization problem of Wiener model is much simpler [5], [6]. For several decades, nonlinear Wiener systems have got much attention by the researchers w.r.t. their properties of system identification.

Chen et al. in [7] explained two identification methods for nonlinear output-error systems excited by dual-rate sampled data. Hagen et al. in [8] proposed a method based on maximum likelihood estimation to get Wiener-model parameters and has tested the effectiveness of the deduced algorithm.

An identification problem of Wiener models with invertible nonlinear function has raised the concern in recent past [9]. For example, in [10], a Hammerstein-Wiener model with invertible nonlinear parts has been identified using a blind approach. In the research paper [11], an identification method for FIR Wiener systems has been investigated. This method considered the polynomial nonlinearity to be non-invertible with unknown parameters and has given three different methods for estimating unknown non-linearity. Janczak in [12] has reviewed many approaches to identify Wiener and Hammerstein models applicable for the neural network and polynomial models.

There are certain uncertainties encountered during the modeling of nonlinear systems. These uncertainties include unknown parameters and undefined nonlinearity. In order to overcome these uncertainties, data-based models are required [5], [13]. In this article, LMS algorithm and its variants named leaky LMS and modified leaky LMS algorithms are employed to overcome the above-mentioned uncertainties. Since LMS based algorithms are very much suitable for the real-time applications, it can estimate parameters even when all the data are not available but come at every instant. The favorable features of LMS based algorithms are its simplicity and easy to implement. The main difficulty with basic LMS algorithm is the weight drift problem which may occur in non-ideal situations [14], [15]. In order to combat the weight drifting problem, a leaky LMS algorithm is derived where a leakage factor is introduced to control the weight update [16]. In order to enhance the rate of convergence of leaky LMS, modified leaky LMS is taken into consideration. The static nonlinear function is assumed to be invertible to justify intermediate variables. The performance of the deduced methodology is determined using the mean-square deviation (MSD), the excess mean-square error (EMSE) and mean square error (MSE) plots.

The general notations used in this article are as follows: the capital letter with the bold case is used for a matrix, small letter with bar is used for column vector and the letter with neither bar nor bold case is used to represent a scalar quantity. $(\cdot)^T$ denotes the transpose operation.

Following the introduction in Section I, Wiener system and its corresponding model identification problem is formulated in Section II. LMS algorithm, leaky LMS algorithm and modified leaky LMS algorithm for identifying parameters of Wiener nonlinear model are explained in Sections III, IV

and V respectively. In Section VI, an example supporting the effectiveness of the deduced methodology is described and simulation results are obtained. Finally, the conclusion of the article is provided in Section VII.

II. WIENER SYSTEM DESCRIPTION AND MODEL IDENTIFICATION FORMULATION

A Wiener nonlinear system comprises of an LTI system in series with a static nonlinear function $F(\cdot)$. A Wiener nonlinear system can be expressed as shown in Fig. 1.

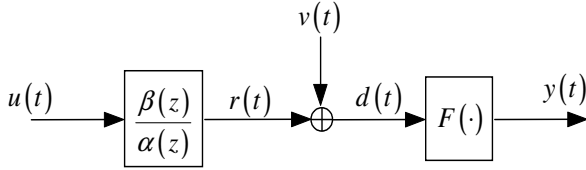


Fig. 1: Wiener nonlinear system [17]

$\frac{\beta(z)}{\alpha(z)}$ represents the traditional transfer function of the LTI system, $u(t)$ and $y(t)$ denote the input and output of the Wiener system respectively. The output of LTI system is denoted by $r(t)$ and can be expressed as

$$r(t) = \frac{\beta(z)}{\alpha(z)}u(t). \quad (1)$$

The Wiener nonlinear system is indebted by a stochastic white Gaussian noise $v(t)$ with mean zero and variance σ^2 . From Fig.1, the output of Wiener system can be expressed as

$$y(t) = F(d(t)), \quad (2)$$

where the intermediate variable $d(t)$ can be written as

$$d(t) = r(t) + v(t) = \frac{\beta(z)}{\alpha(z)}u(t) + v(t). \quad (3)$$

Let the LTI block have a rational transfer function with numerator and denominator are represented by polynomials having backward time shift operator z^{-1} :

$$\begin{aligned} \alpha(z) &= 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots + \alpha_{n_\alpha} z^{-n_\alpha} \\ \beta(z) &= \beta_1 z^{-1} + \beta_2 z^{-2} + \dots + \beta_{n_\beta} z^{-n_\beta}. \end{aligned} \quad (4)$$

The orders n_α and n_β are supposed to be previously known, and static nonlinear function $F(\cdot)$ is assumed to be invertible. The relationship between intermediate variable $d(t)$ and output of the nonlinear system $y(t)$ can be expressed as

$$d(t) = F^{-1}(y(t)) = \sum_{k=1}^m c_k f_k(y(t)), \quad (5)$$

where $f_k(\cdot)$ are basis functions for nonlinearity of the system. These basis functions can be polynomials, radials and splines. The order of nonlinearity m is considered to be foreknown.

Using equation (4) in (3), $d(t)$ can be explicitly written as

$$\begin{aligned} d(t) &= \sum_{i=1}^{n_\alpha} \alpha_i [v(t-i) - d(t-i)] \\ &\quad + \sum_{j=1}^{n_\beta} \beta_j u(t-j) + v(t). \end{aligned} \quad (6)$$

From equation (5) and (6), we get

$$\begin{aligned} \sum_{k=1}^m c_k f_k(y(t)) &= \sum_{i=1}^{n_\alpha} \alpha_i [v(t-i) - d(t-i)] \\ &\quad + \sum_{j=1}^{n_\beta} \beta_j u(t-j) + v(t). \end{aligned} \quad (7)$$

Without losing any generality, it can be assumed that $c_1 = 1$. Hence (7) can be rewritten as

$$\begin{aligned} f_1(y(t)) &= \sum_{i=1}^{n_\alpha} \alpha_i [v(t-i) - d(t-i)] \\ &\quad + \sum_{j=1}^{n_\beta} \beta_j u(t-j) - \sum_{k=2}^m c_k f_k(y(t)) + v(t). \end{aligned} \quad (8)$$

Typically, the polynomial representation is chosen as the nonlinear basis functions because of its simple implementation and easy to analyze. Equation (8) is in the linear regression form and can be written in a simplified way for polynomial nonlinear basis functions as shown below:

$$y(t) = \bar{\phi}^T(t) \bar{\theta} + v(t) \quad (9)$$

where

$$\bar{\theta} = [\bar{\theta}_1^T, c_2, \dots, c_m]^T \in \mathbb{R}^{n=n_\alpha+n_\beta+m-1} \quad (10)$$

with

$$\bar{\theta}_1 = [\alpha_1, \dots, \alpha_{n_\alpha}, \beta_1, \dots, \beta_{n_\beta}]^T \in \mathbb{R}^{n_1=n_\alpha+n_\beta} \quad (11)$$

and

$$\bar{\phi}(t) = [\bar{\phi}_1^T(t), -y^2(t), \dots, -y^m(t)] \in \mathbb{R}^n \quad (12)$$

with

$$\bar{\phi}_1(t) = \begin{bmatrix} v(t-1) - d(t-1), \dots, v(t-n_\alpha) \\ -d(t-n_\alpha), u(t-1), \dots, u(t-n_\beta) \end{bmatrix} \in \mathbb{R}^{n_1}. \quad (13)$$

From equation (6), (11) and (13), the intermediate unknown variable $d(t)$ can be expressed as

$$d(t) = \bar{\phi}_1^T(t) \bar{\theta}_1 + v(t). \quad (14)$$

The aim of this paper is to estimate parameter vectors $\bar{\theta}$ and $\bar{\theta}_1$ using LMS algorithm.

III. LEAST MEAN SQUARE ALGORITHM

In this section, we design an LMS based Wiener identification methodology to estimate model parameter vector $\bar{\theta}$. The

quadratic cost function for the Wiener model is defined as

$$J(\bar{\theta}) = \sum_{t=1}^L |y(t) - \bar{\phi}^T(t)\bar{\theta}|^2, \quad (15)$$

here the data length is considered to very large than n i.e., $L \gg n$.

Let us define the estimate of $\bar{\theta}$ at any time t as $\hat{\theta} = [\hat{\theta}_1^T, \hat{c}_2, \dots, \hat{c}_m]^T$. The aim is to minimize cost function (15) w.r.t. $\bar{\theta}$ and get the model parameter estimates. Using LMS minimization, the update equation for estimating $\bar{\theta}$ is obtained as

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \mu \bar{\phi}(t) e(t), \quad (16)$$

where $\mu > 0$ is a small positive constant step-size involved while minimizing the quadratic cost function (15) using LMS algorithm,

$$e(t) = y(t) - \bar{\phi}^T(t)\hat{\theta}(t), \quad (17)$$

$$\bar{\phi}(t) = [\bar{\phi}_1^T(t), -y^2(t), \dots, -y^m(t)], \quad (18)$$

with

$$\bar{\phi}_1(t) = \begin{bmatrix} v(t-1) - d(t-1), \dots, v(t-n_\alpha) \\ -d(t-n_\alpha), u(t-1), \dots, u(t-n_\beta) \end{bmatrix}. \quad (19)$$

In this article, a constant step-size is considered. It is also possible to use a step-size of decaying nature i.e., time dependent step-size $\mu(t) \geq 0$. The decaying step-size $\mu(t)$ should satisfy the following two conditions,

$$\sum_{t=0}^{\infty} \mu(t) = \infty, \quad \lim_{t \rightarrow \infty} \mu(t) = 0. \quad (20)$$

However, the rate of convergence will be slower and as $t \rightarrow \infty$, the step-size will die out and turns off the adaptation [18]. Due to these reasons, we have considered a constant step-size in this article. The stability bound for step-size is given as

$$0 < \mu < \frac{2}{Tr[R_{\hat{\phi}}]} \quad (21)$$

i.e., the step-size should be chosen within the bound otherwise algorithm may diverge and becomes unstable. where $R_{\hat{\phi}}$ is the covariance matrix of $\hat{\phi}(t)$ [19].

Notice that $\bar{\phi}_1(t)$ in (19) contains unmeasured terms $v(t-i)$ and $d(t-i)$, which makes the algorithm infeasible to estimate $\hat{\theta}(t+1)$. To overcome this situation, auxiliary model identification based approach is implied.

Let $\hat{v}(t)$ and $\hat{d}(t)$ are the estimates of $v(t)$ and $d(t)$ respectively. Then, define

$$\hat{\phi}_1(t) = \begin{bmatrix} \hat{v}(t-1) - \hat{d}(t-1), \dots, \hat{v}(t-n_\alpha) \\ -\hat{d}(t-n_\alpha), u(t-1), \dots, u(t-n_\beta) \end{bmatrix} \in \mathbb{R}^{n_1}, \quad (22)$$

$$\hat{\phi}(t) = \begin{bmatrix} \hat{\phi}_1^T(t), -f_2(y(t)), \dots, -f_m(y(t)) \end{bmatrix} \in \mathbb{R}^n. \quad (23)$$

Replacing $\bar{\phi}_1(t)$, $\bar{\phi}(t)$, $\bar{\theta}_1(t)$ and $\bar{\theta}(t)$ in equation (9) and

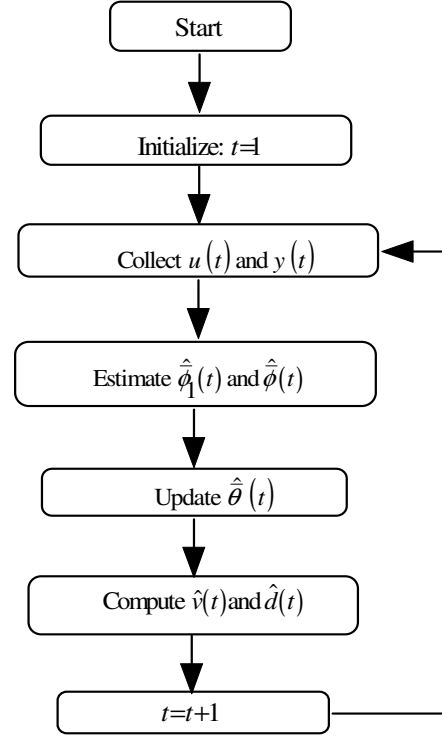


Fig. 2: Flowchart for estimating $\bar{\theta}$ using LMS algorithm

(14) with $\hat{\phi}_1(t)$, $\hat{\phi}(t)$, $\hat{\theta}_1(t)$ and $\hat{\theta}(t)$, we get

$$\begin{aligned} \hat{v}(t) &= y(t) - \hat{\phi}^T(t)\hat{\theta}(t) \\ \hat{d}(t) &= \hat{\phi}_1^T(t)\hat{\theta}_1(t) + \hat{v}(t). \end{aligned} \quad (24)$$

Now the update equation for $\hat{\theta}(t)$ can be written as

$$\hat{\theta}(t+1) = \hat{\theta}(t) + \mu \hat{\phi}(t) \hat{e}(t), \quad (25)$$

where

$$\hat{e}(t) = y(t) - \hat{\phi}^T(t)\hat{\theta}(t) \quad (26)$$

Wiener-LMS algorithm steps to estimate model parameter $\bar{\theta}(t)$ are listed as following:

- 1) Initialize the values $\hat{\theta}(t) = 0$, $u(t) = 0$, $y(t) = 0$, $\hat{d}(t) = 0$, $\hat{v}(t) = 0$ for $t \leq 0$. For $t > 0$
- 2) Input data $u(t)$ and output data $y(t)$ are collected.
- 3) Using (22) and (23), form $\hat{\phi}_1(t)$ and $\hat{\phi}(t)$.
- 4) The estimation of the Wiener modeling parameter $\hat{\theta}(t)$ is updated using (25).
- 5) Find $\hat{\theta}_1(t)$ from $\hat{\theta}(t)$ using $\hat{\theta} = [\hat{\theta}_1^T, \hat{c}_2, \dots, \hat{c}_m]^T$.
- 6) The value of $\hat{v}(t)$ and $\hat{d}(t)$ are calculated using (24).
- 7) If $t < L$, then repeat steps 2-6. If $t = L$ reaches, stop the process and get the estimates of model parameter using (10).

The flowchart presenting the steps involved in Wiener-LMS algorithm for the estimation of parameter vector $\bar{\theta}(t)$ is depicted in Fig. 2.

IV. LEAKY LEAST MEAN SQUARE ALGORITHM

One of the drawback concerning LMS algorithm is the drifting problem which has been analyzed in [15], [16]. This situation made LMS algorithm to generate unbounded estimates for bounded input with large eigen spread. The solution of this problem is to include leakage factor in the cost function which bounds the parameter estimates.

Consider the new instantaneous cost function for leaky-LMS as [14]

$$J(t) = \hat{e}^2(t) + \gamma \hat{\theta}^T(t) \hat{\theta}(t), \quad (27)$$

where $(0 < \gamma < 1)$ is the leakage factor which is user selective. The term $\gamma \hat{\theta}^T(t) \hat{\theta}(t)$ can be viewed as the regularization term. Using the fundamentals of LMS algorithm, the recursive update equation for leaky LMS can be written as

$$\hat{\theta}(t+1) = (1 - \mu\gamma) \hat{\theta}(t) + \mu \hat{\phi}(t) \hat{e}(t). \quad (28)$$

The stability bound of the step size μ is given by

$$0 < \mu < \frac{2}{\gamma + \lambda_{\max}(R_{\hat{\phi}})} \quad (29)$$

or $0 < \mu < \frac{2}{\gamma + Tr[R_{\hat{\phi}}]}$

V. MODIFIED LEAKY LMS

The drawback concerning leaky LMS is its low convergence rate. To enhance the converging rate of leaky LMS, a modified leaky LMS is proposed in [14]. A new cost function is defined where the sum of exponentials are employed in the cost function of leaky LMS. The modified instantaneous cost function is given as [14]

$$J(t) = (\exp(\hat{e}(t)) + \exp(-\hat{e}(t)))^2 + \gamma \hat{\theta}^T(t) \hat{\theta}(t). \quad (30)$$

The recursive update equation for modified leaky LMS is given by

$$\hat{\theta}(t+1) = (1 - \mu\gamma) \hat{\theta}(t) + 2\mu \hat{\phi}(t) \sinh(\hat{e}(t)) \quad (31)$$

VI. EXAMPLE FOR NUMERICAL TEST

Let us consider a Wiener nonlinear system with finite impulse response as the transfer function of LTI system

$$d(t) = \frac{\beta(z)}{\alpha(z)} u(t) + v(t),$$

$$\alpha(z) = 1,$$

$$\beta(z) = \beta_1 z^{-1} + \beta_2 z^{-2} = 0.8z^{-1} + 0.33z^{-2},$$

and the inverse of the nonlinear function

$$d(t) = y(t) - 0.5y^3(t).$$

Persistently excited signal is taken as the input to the system under consideration. The input considered is of zero mean and unit variance. The noise affecting the system is considered to be the white Gaussian noise having mean zero and standard deviation 0.01. Initialization for the variables to

estimate parameter of interest are taken as $\hat{\theta}(t) = 0$, $u(t) = 0$, $y(t) = 0$, $\hat{d}(t) = 0$, $\hat{v}(t) = 0$ for $t = 0$. Applying the different variants of LMS algorithm to obtain the Wiener model parameter vector $\hat{\theta}(t)$.

Let us define some performance indexes to measure the effectiveness of the proposed algorithm; mean-square deviation (MSD) at time iteration t as:

$$MSD_t = E \left\| \hat{\theta}(t) - \bar{\theta} \right\|^2 = E \left\| \tilde{\theta}(t) \right\|^2, \quad (32)$$

where $\tilde{\theta} = \hat{\theta}(t) - \bar{\theta}$. Excess-mean square error (EMSE) at time iteration t

$$EMSE_t = E \left| \hat{\phi}^T(t) \tilde{\theta}(t) \right|^2. \quad (33)$$

Mean square error (MSE) at time iteration t

$$MSE_t = E |e(t)|^2, \quad (34)$$

where $\hat{e}(t)$ is calculated from (26). Here the expectation operator is introduced for taking the average of the estimates while a number of times the algorithm runs.

The steady-state values for these performance indexes can be calculated as:

$$MSD = \lim_{t \rightarrow \infty} E \left\| \hat{\theta}(t) - \theta \right\|^2 = \lim_{t \rightarrow \infty} E \left\| \tilde{\theta}(t) \right\|^2$$

$$EMSE = \lim_{t \rightarrow \infty} E \left| \hat{\phi}^T(t) \tilde{\theta}(t-1) \right|^2 \quad (35)$$

$$MSE = \lim_{t \rightarrow \infty} E |e(t)|^2$$

The performance of the deduced algorithm is analyzed with the help of learning curves; Mean Square Deviation (MSD), Excess Mean Square Error (EMSE) and Mean Square Error (MSE) plots with noise variance $\sigma^2 = 0.01^2$. A constant step-size of 0.01 and leakage factor $\gamma = 0.0001$ is considered for updating model parameter vector. A total of 250 experiments is

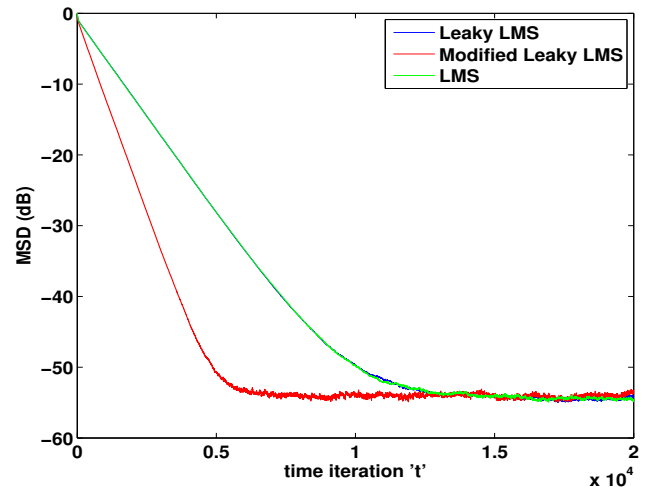


Fig. 3: MSD for estimating $\bar{\theta}$ using LMS and its different variant algorithms

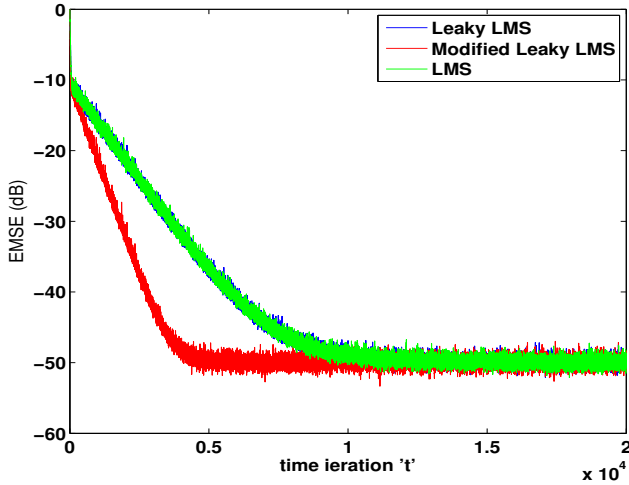


Fig. 4: EMSE for estimating $\bar{\theta}$ using LMS and its different variant algorithms

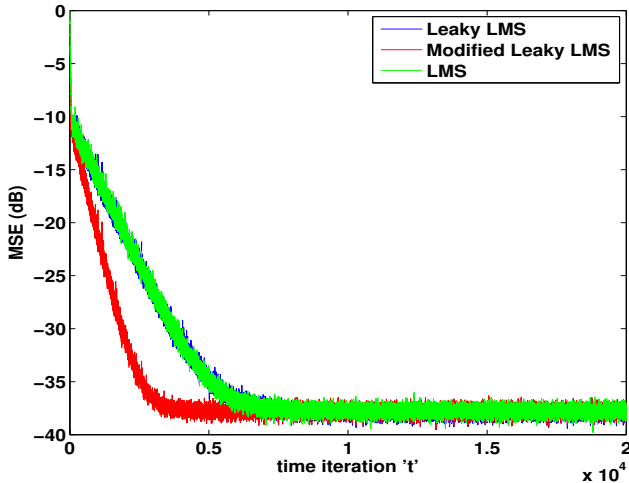


Fig. 5: MSE for estimating $\bar{\theta}$ using LMS and its different variant algorithms

considered and each experiment runs for 20000 time iterations. The curves are plotted by taking the performance average of all the experiments. The learning curves obtained are shown in Fig. (3), Fig. (4) and Fig. (5). Plots show that modified leaky LMS have good convergence rate when compared to other two variants as explained in Section V. The learning curves also called performance plots, have very low steady state values, hence the algorithms designed are very efficient in estimating Wiener model parameters.

VII. CONCLUSION

This article presented the Wiener model parameter estimation using LMS algorithm with its two modified variants. The methodology is assumed to pursue the invertible nonlinear function which is justified theoretically. The simulation

results obtained using MATLAB software demonstrate that the proposed methodologies are very efficient and work well in presence of noise. The recommended algorithms can be broadened to identify the parameters of other nonlinearity based systems with different types of noise in them.

REFERENCES

- [1] F. Ding, "System identification-performances analysis for identification methods," *Science Press*, 2014.
- [2] H. Li, Y. Gao, P. Shi, and H.-K. Lam, "Observer-based fault detection for nonlinear systems with sensor fault and limited communication capacity," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2745–2751, 2016.
- [3] X. Luan, P. Shi, and F. Liu, "Stabilization of networked control systems with random delays," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 4323–4330, 2011.
- [4] S. A. Ghauri and M. F. Sohail, "System identification using LMS, NLMS and RLS," in *Research and Development (SCORED), 2013 IEEE Student Conference on*. IEEE, 2013, pp. 65–69.
- [5] H.-X. Li and C. Qi, *Spatio-temporal modeling of nonlinear distributed parameter systems: A time/space separation based approach*. Springer Science & Business Media, 2011, vol. 50.
- [6] J. C. Gomez and E. Baeyens, "Identification of block-oriented nonlinear systems using orthonormal bases," *Journal of Process Control*, vol. 14, no. 6, pp. 685–697, 2004.
- [7] J. Chen and R. Ding, "Two identification methods for dual-rate sampled-data nonlinear output-error systems," *Mathematical Problems in Engineering*, vol. 2014, 2014.
- [8] A. Hagenblad, L. Ljung, and A. Wills, "Maximum likelihood identification of Wiener models," *Automatica*, vol. 44, no. 11, pp. 2697–2705, 2008.
- [9] A. Hasanov and V. G. Romanov, "An inverse coefficient problem related to elastic-plastic torsion of a circular cross-section bar," *Applied Mathematics Letters*, vol. 26, no. 5, pp. 533–538, 2013.
- [10] E.-W. Bai, "A blind approach to the Hammerstein-Wiener model identification," in *Block-oriented Nonlinear System Identification*. Springer, 2010, pp. 313–332.
- [11] S. L. Lacy and D. S. Bernstein, "Identification of FIR Wiener systems with unknown, non-invertible, polynomial non-linearities," *International Journal of Control*, vol. 76, no. 15, pp. 1500–1507, 2003.
- [12] A. Janczak, *Identification of nonlinear systems using neural networks and polynomial models: a block-oriented approach*. Springer Science & Business Media, 2004, vol. 310.
- [13] H.-X. Li and C. Qi, "Modeling of distributed parameter systems for applications-A synthesized review from time-space separation," *Journal of Process Control*, vol. 20, no. 8, pp. 891–901, 2010.
- [14] T. R. Gwadabe, M. S. Salman, and H. Abuhilal, "A modified Leaky-LMS algorithm," *International Journal of Computer and Electrical Engineering*, vol. 6, no. 3, p. 222, 2014.
- [15] M. Sowjanya, A. Sahoo, and S. Kumar, "Distributed incremental leaky LMS," in *Communications and Signal Processing (ICCSP), 2015 International Conference on*. IEEE, 2015, pp. 1753–1757.
- [16] V. H. Nascimento and A. H. Sayed, "Unbiased and stable leakage-based adaptive filters," *IEEE Transactions on Signal Processing*, vol. 47, no. 12, pp. 3261–3276, 1999.
- [17] A. Hagenblad, *Aspects of the identification of Wiener models*. Division of Automatic Control, Department of Electrical Engineering, Linköping universitet, 1999.
- [18] A. H. Sayed *et al.*, "Adaptation, learning, and optimization over networks," *Foundations and Trends® in Machine Learning*, vol. 7, no. 4-5, pp. 311–801, 2014.
- [19] S. Haykin and B. Widrow, *Least-mean-square adaptive filters*. John Wiley & Sons, 2003, vol. 31.