Applied Soft Computing 13 (2013) 3324–3334

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

NARX model based nonlinear dynamic system identification using low complexity neural networks and robust H_∞ filter

H.K. Sahoo [a,*], P.K. Dash [b], N.P. Rath [c]

[a] IIIT, Bhubaneswar, India
[b] S.O.A. University, Bhubaneswar, India
[c] VSSUT, Burla, India

ARTICLE INFO

Article history:
Received 13 June 2012
Received in revised form 9 December 2012
Accepted 6 February 2013
Available online 15 March 2013

Keywords:
H_∞ filter
EKF
FFRLS
NARX model
Chebyshev polynomials
Legendre polynomials
Wavelet neural network

ABSTRACT

This paper proposes NARX (nonlinear autoregressive model with exogenous input) model structures with functional expansion of input patterns by using low complexity ANN (artificial neural network) for nonlinear system identification. Chebyshev polynomials, Legendre polynomials, trigonometric expansions using sine and cosine functions as well as wavelet basis functions are used for the functional expansion of input patterns. The past input and output samples are modeled as a nonlinear NARX process and robust H_∞ filter is proposed as the learning algorithm for the neural network to identify the unknown plants. H_∞ filtering approach is based on the state space modeling of model parameters and evaluation of Jacobian matrices. This approach is the robustification of Kalman filter which exhibits robust characteristics and fast convergence properties. Comparison results for different nonlinear dynamic plants with forgetting factor recursive least square (FFRLS) and extended Kalman filter (EKF) algorithms demonstrate the effectiveness of the proposed approach.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

In modern applications, devices and systems are increasingly used in regions where they cannot be accurately described with a linear model anymore. This drives a growing need for models and modeling techniques able to adequately describe these systems' behavior. The reason for this great interest in this field is related to the nonlinear nature of the physical phenomenon occurring in the real world, making the linear hypothesis just an approximation of the real behavior. Nonlinear system identification has been used in many applications and these include control, instrumentation, power systems, communication systems, nonlinear amplifiers in the transmission stage, signal processing, neuroscience, and in satellite communications, etc. Estimating nonlinear systems is a very broad problem as it is impossible to propose a structure able to describe efficiently every possible nonlinear system. Hence, the scope is often reduced to focus on nonlinear model structures built with static nonlinearities and linear dynamics. Hammerstein model may be a good approximation for nonlinear plants. They are composed of a static nonlinear gain and linear dynamics. Hammerstein-model based identification [1–4] has been dealt with using both parametric and nonparametric methods. In the parametric approach, the nonlinear element is usually modeled by a finite-order polynomial with unknown coefficients (but known order). Based on this assumption, parameter estimation can be dealt with using correlation techniques using Volterra kernel, recursive algorithms or combinations of these. In case of expectation maximization (EM) [5], a maximum likelihood approach that iteratively maximizes the likelihood of the observations by maximizing the joint likelihood function in each iteration is used. EM algorithms are implemented by iteratively finding the expected value of the joint likelihood function in the first step and maximizing it in the second step. The significance and difficulty of estimating nonlinear systems are widely recognized [6,7]. As a result, there is very large and active research effort directed toward the problem. A key aspect of this activity is that it generally focuses on specific system classes such as those described by neural networks [8–10], nonlinear ARMAX (NARMAX) [11–13,31] to name just some examples.

Evolutionary method for system modeling and time series prediction is also being presented [14–16] which combines the effectiveness of adaptive multi model partitioning filters and GAs' robustness. Selective recursive kernel learning (SRKL) is being proposed [17] with NARX form for online identification of MIMO system. Type-2 neuro-fuzzy system structure [18] is also proposed for identification of dynamic time-varying plants and equalization of time-varying channels. Single layer FLANN model [19,20] based

* Corresponding author. Tel.: +91 9437086263; fax: +91 6743016009.
E-mail addresses: harish@iiit-bh.ac.in (H.K. Sahoo), pkdash.india@gmail.com (P.K. Dash), n_p_rath@hotmail.com (N.P. Rath).

1568-4946/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2013.02.007

identification is quite popular due to its simpler structure requires less computation as compared to multilayer perceptrons (MLP). Chebyshev neural networks (CNN) as well as Legendre polynomial based ANN have powerful representation capabilities whose input is generated by using a subset of Chebyshev polynomials [21,22] and Legendre polynomials [23]. It has already been established that pattern classification and system identification using CNN requires less computation as compared to MLP. Local linear wavelet neural network (LLWNN) model is also being proposed [24] for energy market price forecasting which can also be used for system identification by using scale and dilation of chosen mother wavelet.

In this paper, a nonlinear filter known as H_∞ filter [25–27] has been proposed for identifying the unknown plant parameters considering different challenging test plants in presence white Gaussian noise. The unknown plant is modeled as a NARX process and parameters of this model are obtained using functional expansion capabilities of single layer neural network with the help of Chebyshev, Legendre polynomials as well as trigonometric functions like sine and cosine functions. Also local linear wavelet neural network (LLWNN) model is used with wavelet basis functions for functional expansion of input patterns. The main purpose of choosing H_∞ filter for estimation is due to its robustness to the nature of variation of measurement noise and tracking capability in case of nonlinear and dynamic system under low SNR condition.

This paper is organized as follows. The problem formulation is presented in Section 2. Different types of NARX modeling structures used for system identification are presented in Section 3. In Section 4, different ANN structures are presented with polynomial, trigonometric and wavelet basis expansions of process input and output samples of different NARX models. Proposed H_∞ filtering method for weight updation of neural network is presented in Section 5. The robustness of the proposed filtering algorithm is proved mathematically in Section 6. Simulation results related to the identifications of nonlinear dynamic plants using different input excitations are presented in Section 7. Conclusion is given in Section 8.

2. Problem formulation

This paper considers the problem of identifying the parameters θ for certain members of the following nonlinear state-space model structure [4,25]

$$x_{k+1} = f_k(x_k, u_k, w_k, \theta) \tag{1}$$

$$y_k = h_k(x_k, u_k, v_k, \theta) \tag{2}$$

Here x_k denotes state variable with u_k and y_k denote observed input and output responses. Furthermore θ is a vector of unknown parameters that specifies the mappings $f_k(.)$ and $h_k(.)$ which may be nonlinear and time varying. Where state transition matrix is obtained as $F_k = (\partial f_k/\partial x)|_{x=\hat{x}_k}$ from Eq. (1) and the measurement matrix is obtained as $H_k = (\partial h_k/\partial x)|_{x=\hat{x}_k}$ from measurement Eq. (2) by using Taylor's series expansion. w_k is the process or modeling noise due to modeling, v_k is the measurement noise In such a situation, after linearization using partial derivatives we would like to estimate an arbitrary linear combination of the state x_k, say $z_k = H_k x_k$ using observation y_k. The problem addressed here is the formation of an estimate $\hat{\theta}$ of the parameter vector θ based on N measurements $U_N = [u_1, u_2, \ldots, u_N]$, $Y_N = [y_1, y_2, \ldots, y_N]$ of observed system input and output responses. The method for system identification of a time invariant, causal, discrete time plant is depicted in Figs. 2 and 3. The plant is excited by a signal $u(k)$ and the output $y(k)$ is measured. The plant is assumed to be stable with known parameterization but with unknown value of parameters. The objective is to construct a suitable identification model which when subjected to the same input $u(k)$ as the plant, produces an output $\hat{y}(k)$ in the sense described by $\left\|y - \hat{y}\right\| \le \varepsilon$ for some desired $\varepsilon > 0$ and a suitably defined norm. There are several papers published [28] to explain the concepts of neural network modeling and implementation of nonlinear system identification methods which are extensively used in the area of communication, signal processing, biomedical engineering, etc.

3. Nonlinear input–output models for feed forward neural networks

ARX (Auto Regressive with exogenous inputs) model structure is capable of describing any linear system corrupted by additive noise. One reason to consider more elaborate model such as ARMAX (Auto Regressive Moving Average with Exogenous inputs) models is that same fit can be obtained with fewer parameters. Using same regressor as for the linear models were obtained as NARX, NARMAX (the character N in the model names denotes Nonlinear). The model family based on different noise assumptions should be seen as reasonable candidate models rather than reflecting the true noise descriptions. The regression vector of NARX models [17,29] consists of a finite number of past values of process inputs and outputs:

$$\begin{aligned}\varphi(k) &= [\varphi_y(k), \varphi_u(k)] \\ &= [y(k-1), y(k-2), \ldots, y(k-na), u(k-1), u(k-2), \ldots, u(k-nb)]^T\end{aligned} \tag{3}$$

The most general NARX model is obtained when the whole regression vector is fed into a general nonlinear mapping (NARX1 model in Fig. 1):

$$\hat{y}(k) = f_N[\varphi_y(k), \varphi_u(k)] \tag{4}$$

The model structure is obtained in case of NARX2 model is nonlinear only in the past values of process inputs, while it is linear in the past values of the process outputs.

$$\hat{y}(k) = [I - A(q)]y(k) + f_{N2}[\varphi_u(k)] \tag{5}$$

where q-shift operator $qy(k) = y(k+1)$; $q^{-1}y(k) = y(k-1)$ and $A(q) = 1 + a_1 q^{-1} + \cdots + a_{na} q^{-na}$.

This model displays the very convenient feature that its stability can easily be determined. As a result, it is well suited as the first candidate model within the nonlinear black-box class of models. If the NRX2 model is not flexible enough to represent the considered process, it is reasonable to try a model which is obtained with the assumption that system noise is additive but with variable variance (NARX3 model).

The model structure is obtained as:

$$\hat{y}(k) = f_{N1}[\varphi_y(k)] + f_{N2}[\varphi_u(k)] \tag{6}$$

The above mentioned NARX model structures can be used in case of neural network by functionally expanding the input patterns.

4. ANN structures with polynomial, trigonometric and wavelet expansion

4.1. Chebyshev neural network (CNN)

CNN is a functional link network [21,22] based on Chebyshev polynomials. Chebyshev series is frequently used for approximations to functions and much more efficient than other power series of same degree. Among orthogonal polynomials, the Chebyshev polynomials occupy an important place, expansions in Chebyshev polynomials converge more rapidly than expansion in any other

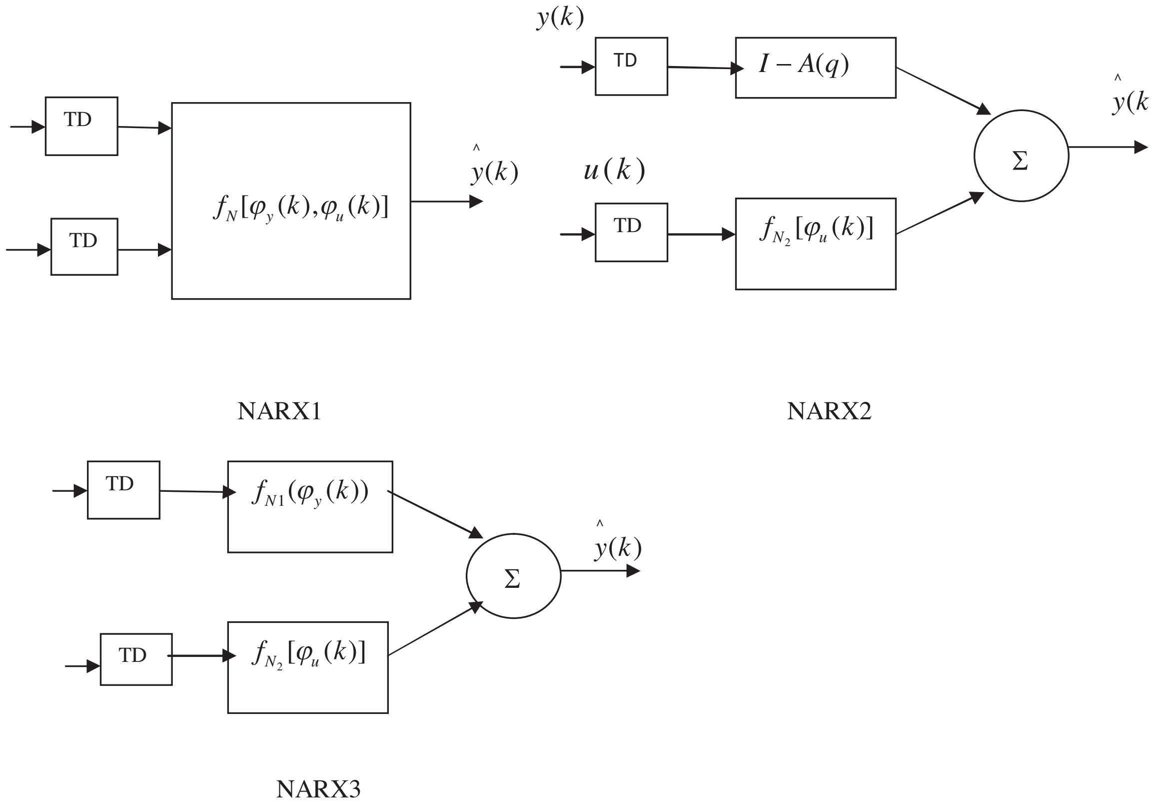

Fig. 1. Model structures from NARX models family (TD-time delay line).

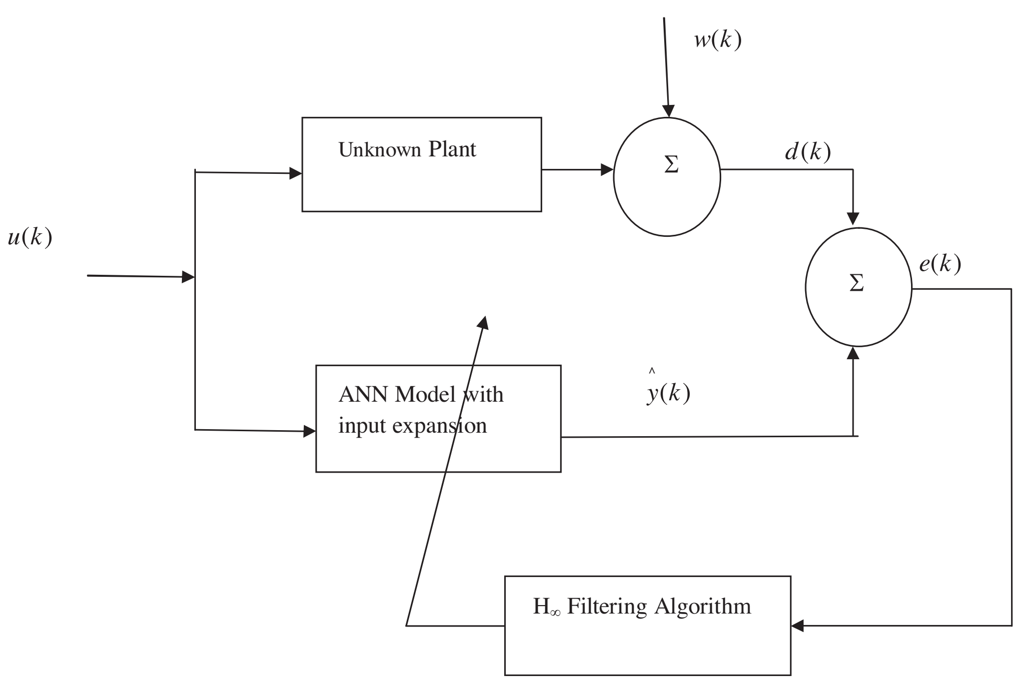

Fig. 2. Nonlinear system identification scheme.

set of polynomials. Hence, in this paper Chebyshev polynomials are considered as basis functions for neural network.

The Chebyshev polynomials can be generated by the following recursive formula:

$$T_{i+1}(x) = 2xT_i(x) - T_{i-1}(x),\ T_0(x) = 1 \tag{7}$$

Consider a two dimensional input pattern $X = [x_1 x_2]^T$. An expanded pattern obtained using Chebyshev functions is given by:

$$\phi = [1\ T_1(x_1)\ T_2(x_1)\ldots T_1(x_2)\ T_2(x_2)\ldots] \tag{8}$$

where $T_i(x_j)$ is Chebyshev polynomials. The different choices of $T_1(x)$ are x, $2x$, $2x - 1$ *and* $2x + 1$. In this paper, $T_1(x)$ is chosen as x.

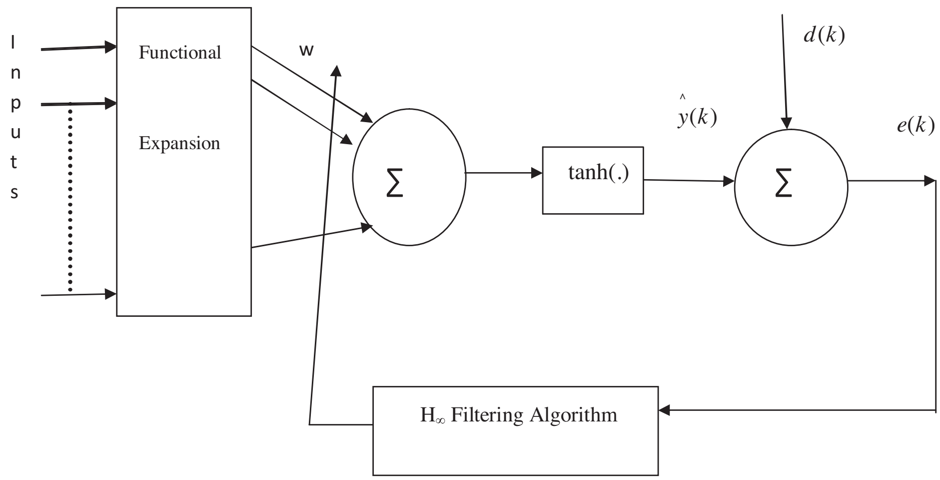

Fig. 3. ANN model with functional expansion and tanh(.) activation function.

4.2. Legendre polynomial based neural network

In this type of neural network the input pattern is expanded into a nonlinear high dimensional space using Legendre polynomials [23] and single layer perceptron network. The Legendre polynomials are denoted as $P_n(x)$ where $n = 0, 1, 2\ldots$ is the order of the polynomial and $-1 \le x \le 1$. These are a set of orthogonal polynomials defined as a solution to the differential equation:

$$\frac{d}{dx}[(1-x^2)\frac{d}{dx}P_n(x)] + n(n+1)P_n(x) = 0 \tag{9}$$

The first few Legendre polynomials are as follows:

$$P_0(x) = 1$$

$$P_1(x) = x$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1)$$

$$P_3(x) = \frac{1}{2}(5x^3 - 3x)$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$$

The higher order polynomials can be derived from the following recursion formula:

$$P_{n+1}(x) = \frac{1}{n+1}[(2n+1)xP_n(x) - nP_{n-1}(x)] \tag{10}$$

4.3. ANN using trigonometric expansion

In this type of ANN structure each feed forward input sample $x(n)$ is nonlinearly expanded using trigonometric expansion scheme [19,20,30] as:

$$\begin{aligned}&[x(n),\ \sin(2\pi x(n)),\ \cos(2\pi x(n)),\ \sin(3.2\pi x(n)),\\ &\cos(3.2\pi x(n))\ldots\sin((2s-1).2\pi x(n)),\ \cos((2s-1).2\pi x(n))]\end{aligned} \tag{11}$$

The symbol s represents the number of sine and cosine expansions of each input sample.

4.4. Local linear wavelet neural network model (LLWNN)

The difference of a local linear wavelet neural network (LLWNN) with conventional wavelet neural network (WNN) is that the connection weights between layers of WNN are replaced by a local linear model [24]. Nonlinear wavelet basis functions are localized in both time space and frequency space. The standard output of WNN is given by

$$f(x) = \sum_{i=1}^{M} w_i\psi_i(x) = \sum_{i=1}^{M} w_i\left|a_i\right|^{-(1/2)}\psi_i\left(\frac{x-b_i}{a_i}\right) \tag{12}$$

where a_i and b_i are scale and translation parameters respectively.

In our model the output of the neural network is given by:

$$y = \sum_{i=1}^{M} w_i\psi_i(x) \tag{13}$$

The mother wavelet is chosen for nonlinear system identification as

$$\psi(x) = \frac{-x^2}{2}e^{-x^2/\sigma^2} \tag{14}$$

where $x = \sqrt{p_1^2 + p_2^2 + \cdots + p_n^2}$ and $p_1, p_2, \ldots, p_n$ form the input set. The error function is defined as $E = [d - w_1\psi_1(x) - w_2\psi_2(x)\ldots - w_n\psi_n(x)]$ and d is the desired output.

5. Robust learning algorithm

The H_∞ filtering algorithm [25–27] is proposed here as robust learning algorithm for weight updation and can be presented mathematically through state prediction and state filter. Through state filtering state variables or weights are updated in an iterative manner such that estimation error will converge to a minimum value.

5.1. State prediction

$$\tilde{x}_{k+1|k} = f(\hat{x}_{k|k}) \tag{15}$$

$$\begin{aligned}P_{k+1|k} &= F_kP_{k|k-1}F_k^{*T} + G_kG_k^{*T}\\ &\quad - F_kP_{k|k-1}[H_k^{*T}L_k^{*T}]R_k^{-1}\begin{bmatrix}H_k\\ L_k\end{bmatrix}P_{k|k-1}F_k^{*T} + Q_k\end{aligned} \tag{16}$$

$$R_k = R_0 + \begin{bmatrix}H_k\\ L_k\end{bmatrix}P_{k|k-1}[H_k^{*T}\ L_k^{*T}] \tag{17}$$

For a given $\gamma_f > 0$,

$$R_0 = \begin{bmatrix} I & 0\\ 0 & -\gamma_f^2 I\end{bmatrix}$$

In Eq. (15), the symbols ∼ and ˆstand for the estimated and predicted values, respectively. For existence of the complex H$_\infty$ filter, $\hat{P}_{k|k-1}$ needs to satisfy the following equation:

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + [H_k^{*T} L_k^{*T}] \begin{bmatrix} I & 0 \\ 0 & -\gamma_f^{-2} I \end{bmatrix} \begin{bmatrix} H_k \\ L_k \end{bmatrix}$$

$$= P_{k|k-1}^{-1} H_k^{*T} H_k - \gamma_f^{-2} L_k^{*T} L_k > 0 \tag{18}$$

In case of H$_\infty$ filter applied to a nonlinear system, the modified equations will be

$$P_{k+1|k} = F_k P_{k|k-1} F_k^{*T} + \beta^2 G_k G_k^T - F_k P_{k|k-1} [H_k^{*T}\ H_k^{*T}] R_k^{-1} \begin{bmatrix} H_k \\ L_k \end{bmatrix} P_{k|k-1} F_k^{*T} + Q_k \tag{19}$$

$$P_{k|k}^{-1} = P_{k|k-1}^{-1} + (1-\gamma_f^2) H_k^{*T} H_k > 0 \quad (L_k = H_k) \tag{20}$$

5.2. State filter

$$\hat{x}_{k|k} = \tilde{x}_{k|k-1} + K_k (y_k - H_k \tilde{x}_{k|k-1}) \tag{21}$$

The gain K_k is calculated to be

$$K_k = P_{k|k-1} H_k^{*T} (R_k + H_k P_{k|k-1} H_k^{*T})^{-1} \tag{22}$$

where $\hat{x}_{k|k}$ is the state variable after filtering; $\tilde{x}_{k|k-1}$ the state variable after estimation; K_k the gain; $y_k - H\tilde{x}_{k|k-1}$ is the innovation vector.

For existence of the H$_\infty$ filter, $\hat{P}_{k|k-1}$ needs to satisfy the following equations

$$\hat{P}_{k|k}^{-1} = \hat{P}_{k|k-1}^{-1} + [H_k^{*T} L_k^{*T}] \begin{bmatrix} I & 0 \\ 0 & -\gamma_f^2 I \end{bmatrix} \begin{bmatrix} H_k \\ L_k \end{bmatrix}$$

$$= \hat{P}_{k|k-1}^{-1} H_k^{*T} H_k - \gamma_f^{-2} L_k^{*T} L_k > 0 \tag{23}$$

γ_f is a scalar weighting factor chosen to be greater than 1 used in Eq. (17) to update the Measurement error covariance and also in Eq. (19) to update the estimation error covariance as per Riccati recursion formula.

6. Robustness of the proposed filtering algorithm

The robustness of H$_\infty$ filter [32] is proved by showing that MSE (Mean Square Error) due to estimation of neural network weights required to identify the unknown plant is exponentially bounded. If MSE is proved to be bounded, the proposed learning algorithm for neural network weight updation is definitely stable and robust with respect to modeling and measurement error.

For convenience of this proof, one step formulation in terms of the a priori variables are considered. The nonlinear discrete time stochastic system with incomplete information is represented by

$$x_{k+1} = f_k(x_k) + \Delta_{1,k} + w_k \tag{24}$$

$$y_k = h_k(x_k) + \Delta_{2,k} + v_k \tag{25}$$

The variables $\Delta_{1,k}$, $\Delta_{2,k}$ mean uncertainty in dynamic equation and measurement equation due to incomplete process noise covariance, incomplete measurement noise covariance, incomplete model coefficients and unknown measurement bias. $\Delta_{1,k}$, $\Delta_{2,k}$ are independent of w_k, v_k and x_k. If system information is perfectly known, $\Delta_{1,k}$, $\Delta_{2,k}$ are all zero. The functions $f_k(x_k)$ and $h_k(x_k)$ can be expanded by Taylor series expansion as

$$f_k(x_k) = f_k(\hat{x}_k) + F_k(x_k - \hat{x}_k) + \phi_f(x_k, \hat{x}_k) \tag{26}$$

$$h_k(x_k) = h_k(\hat{x}_k) + H_k(x_k - \hat{x}_k) + \phi_h(x_k, \hat{x}_k) \tag{27}$$

where

$$F_k = \frac{\partial f_k}{\partial x}|_{x=\hat{x}_k} \tag{28}$$

$$H_k = \frac{\partial h_k}{\partial x}|_{x=\hat{x}_k} \tag{29}$$

The variables ϕ_f and ϕ_h indicate higher order terms in the expansion. The state update equation can be given by:

$$\hat{x}_{k+1} = f_k(\hat{x}_k) + \bar{K}_k[y_k - h_k(x_k)] \tag{30}$$

Also we can define the estimation error as:

$$\begin{aligned}
\xi_{k+1} &= x_{k+1} - \hat{x}_{k+1} = [f_k(x_k) + \Delta_{1,k} + w_k] \\
&\quad - [f_k(\hat{x}_k) + \bar{k}_k[y_k - h_k(\hat{x}_k)]] \\
&= [f_k(\hat{x}_k) + F_k(x_k - \hat{x}_k) + \phi_f(x_k, \hat{x}_k) + \Delta_{1,k} + w_k] \\
&\quad - [f_k(\hat{x}_k) + \bar{K}_k[h_k(x_k) + \Delta_{2,k} + v_k - h_k(\hat{x}_k)]] \\
&= [F_k(x_k - \hat{x}_k) + \phi_f(x_k, \hat{x}_k) + \Delta_{1,k} + w_k] - \bar{K}_k[h_k(\hat{x}_k) \\
&\quad + H_k(x_k - \hat{x}_k) + \phi_h(x_k, \hat{x}_k) + \Delta_{2,k} + v_k - h_k(\hat{x}_k)]
\end{aligned} \tag{31}$$

$$\xi_{k+1} = (F_k - \bar{K}_k H_k)(x_k - \hat{x}_k) + \phi_f(x_k, \hat{x}_k) - \bar{K}_k \phi_h(x_k, \hat{x}_k) + w_k - \bar{K}_k v_k + \Delta_{1,k} - \bar{K}_k \Delta_{2,k}$$

$$\xi_{k+1} = (F_k - \bar{K}_k H_k)(x_k - \hat{x}_k) + \phi_f(x_k, \hat{x}_k) - \bar{K}_k \phi_h(x_k, \hat{x}_k) + w_k - \bar{K}_k v_k + \Delta_{1,k} - \bar{K}_k \Delta_{2,k}$$

$$\xi_{k+1} = (F_k - \bar{K}_k H_k)\xi_k + r_k + s_k + u_k$$

where

$$r_k = \phi_f(x_k, \hat{x}_k) - \bar{K}_k \phi_h(x_k, \hat{x}_k) \tag{32}$$

$s_k = w_k - \bar{K}_k v_k$ and $u_k = \Delta_{1,k} - \bar{K}_k \Delta_{2,k}$.

The stochastic process ξ_k is said to be exponentially bounded in mean square if there are real numbers $\bar{\eta}$, $\nu > 0$ and $0 < \theta < 1$ such that

$$E[\|\xi_k\|^2] \le \bar{\eta} \|\xi_0\|^2 \theta^k + \nu \tag{33}$$

holds for every $k \ge 0$. Also the stochastic process is said to be bounded with probability one if $\sup_{k\ge 0} \|\xi_k\| < \infty$ holds with probability one.

Condition 1. Let the following assumptions hold.

(a) There are positive real numbers $\bar{a}, \bar{c}, \bar{p}, \underline{p}, \underline{q}, \underline{r} > 0$ such that the following bounds on various matrices hold for every

$$\|F_k\| \le \bar{a}, \le \bar{c}, \underline{p} I_n \le \bar{P}_k \le \bar{p} I_n \tag{34}$$

$$\underline{q} I_n \le Q_k, \underline{r} I_m \le R_k \tag{35}$$

(b) F_k is nonsingular for every $k \ge 0$.

(c) There are positive real numbers $\kappa_{\varphi f}, \kappa_{\varphi h}, \varepsilon_{\varphi f}, \varepsilon_{\varphi h} > 0$ such that nonlinear function ϕ_f and ϕ_h are bounded as:

$$\|\phi_f(x_k, \hat{x}_k)\| \le \kappa_{\varphi f} \|x_k - \hat{x}_k\|^2 \tag{36}$$

$$\|\phi_h(x_k, \hat{x}_k)\| \le \kappa_{\varphi h} \|x_k - \hat{x}_k\|^2 \tag{37}$$

$$\text{for } \|x_k - \hat{x}_k\| \le \varepsilon_{\varphi f} \text{ and } \|x_k - \hat{x}_k\| \le \varepsilon_{\varphi h} \tag{38}$$

(d) There are positive real numbers $\kappa_{\Delta 1}, \kappa_{\Delta 2} > 0$ such that uncertainty terms $\Delta_{1,k}$, $\Delta_{2,k}$ are bounded via

$$\|\Delta_{1,k}\| \le \kappa_{\Delta 1}, \|\Delta_{2,k}\| \le \kappa_{\Delta 2} \tag{39}$$

Lemma 1. *Assume there is a stochastic process $V_k(\xi_k)$ as well as real numbers $\underline{v}$, $\bar{v}$, $\mu > 0$ and $0 < \beta \leq 1$ such that*

$$\underline{v}\left\|\xi_k\right\|^2 \leq V_k(\xi_k) \leq \bar{v}\left\|\xi_k\right\|^2 \tag{40}$$

$$E[V_{k+1}(\xi_{k+1})|\xi_k] - V_k(\xi_k) \leq \mu - \beta V_k(\xi_k) \tag{41}$$

are fulfilled for every solution of (31). Then the stochastic process is exponentially bounded in mean square i.e. we have

$$E[\left\|\xi_k\right\|^2] \leq \frac{\bar{v}}{\underline{v}} E[\left\|\xi_0\right\|^2](1-\beta)^k + \frac{\mu}{\underline{v}} \sum_{i=1}^{k-1} (1-\beta)^i \tag{42}$$

for every $k \geq 0$. Moreover the stochastic process is bounded with probability one.

Lemma 2. *Under condition 1, there is a real number $0 \leq \bar{\chi} \leq 1$ such that satisfies the inequality ($\prod_k = \bar{P}_k^{-1}$)*

$$(F_k - \bar{K}_k H_k)^T \prod_{k+1} (F_k - \bar{K}_k H_k) \leq (1 - \bar{\chi}) \prod_k \tag{43}$$

where $k \geq 0$,

$$\bar{\chi} = \frac{\underline{q}}{[\bar{p}(\bar{a} + \bar{k}\bar{c})^2 + \bar{q}}, \bar{k} = \frac{\lambda_k \bar{a}\bar{p}\bar{c}}{\alpha_k \underline{r}} \tag{44}$$

7. Simulation results for identification of dynamic time varying plants

In this section some experimental simulation results of dynamic nonlinear system identification are presented. All the estimation results are obtained considering the standard test systems in presence of white Gaussian noise with SNR 30 dB. The results of the proposed learning method are being compared with FFRLS and EKF. For H_∞ filter the measurement noise covariance R_k and process noise covariance Q_k are taken as 0.01. For FFRLS algorithm the forgetting factor is taken as 0.99. Same input excitation is given to the actual system and the neural network model. The inputs to the neural network model $[y(k-1), y(k-2), y(k-3), \ldots, u(k), u(k-1), \ldots]$ are expanded using Chebyshev polynomials, Legendre polynomials, and trigonometric expansions using sine and cosine functions as well as wavelet basis function. Then the sum of the estimated weighted inputs is passed through the activation tanh(.) function to generate the estimated output signal. The MSE (Mean

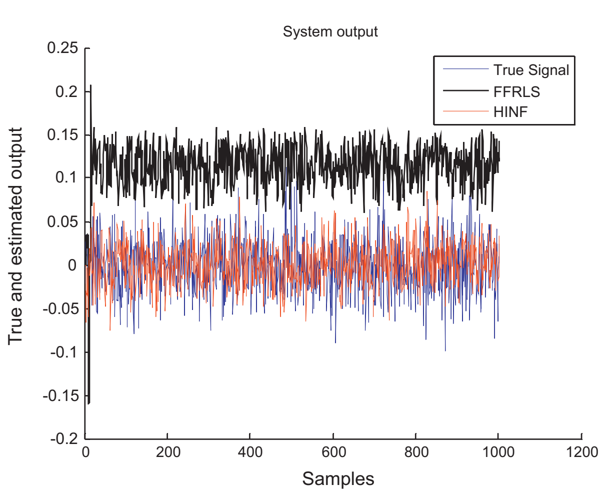

Fig. 4. True and estimated output using trigonometric function (Example-1).

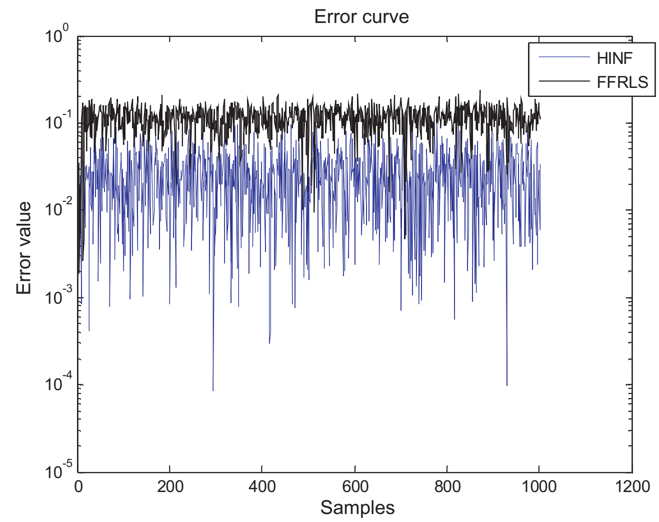

Fig. 5. Absolute error between true and estimated output (Example-1).

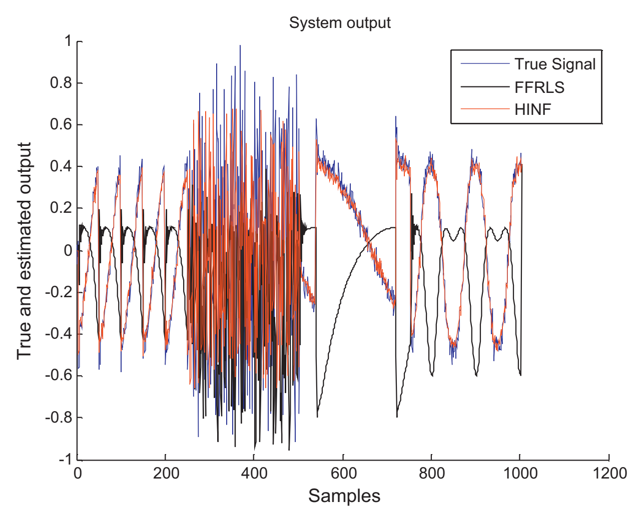

Fig. 6. True and estimated output for input excitation-1 using CNN (Example-2).

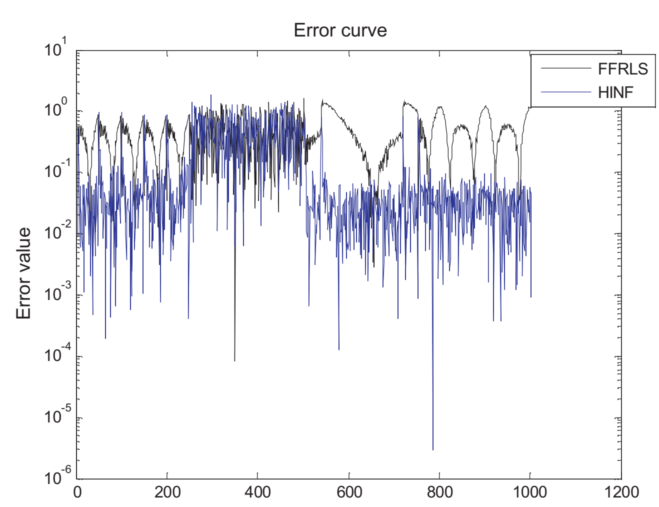

Fig. 7. Absolute error between true and estimated output for input excitation-1 using CNN (Example-2).

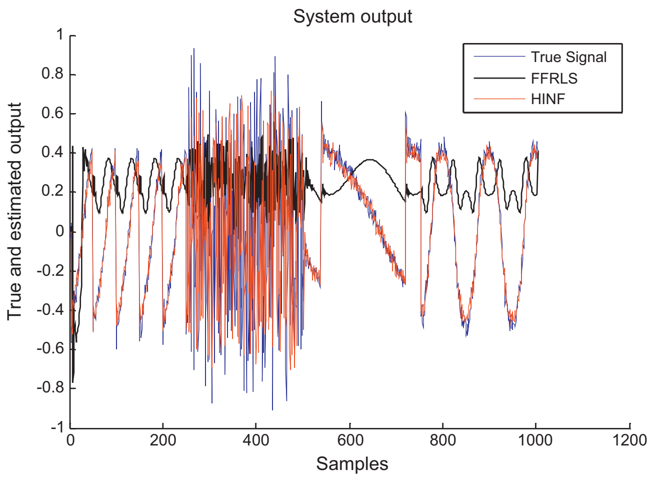

Fig. 8. True and estimated output for input excitation-1 using Legendre polynomial NN (Example-2).

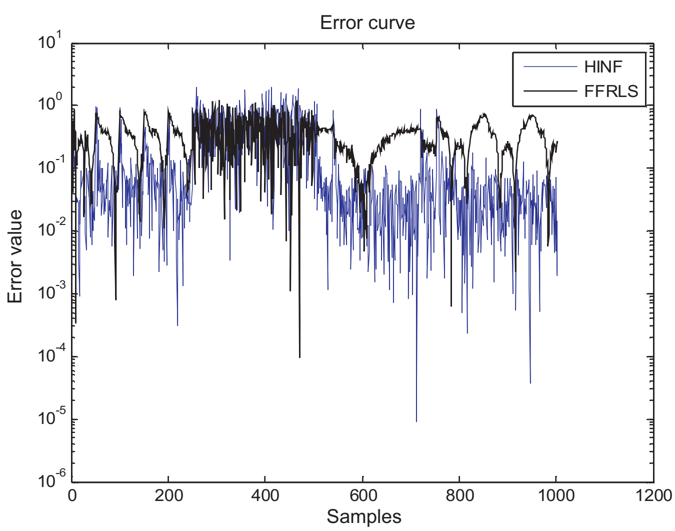

Fig. 9. Absolute error between true and estimated output for input excitation-1 Legendre polynomial NN (Example-2).

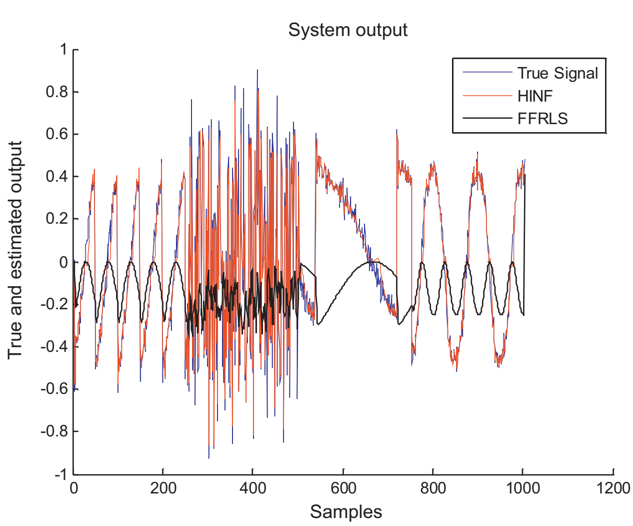

Fig. 10. True and estimated output for input excitation-1 using LLWNN (Example-2).

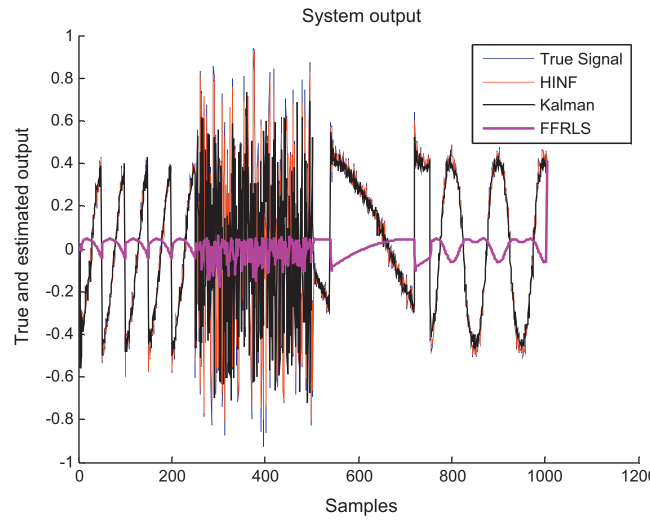

Fig. 11. True and estimated output for input excitation-1 using trigonometric NN (Example-2).

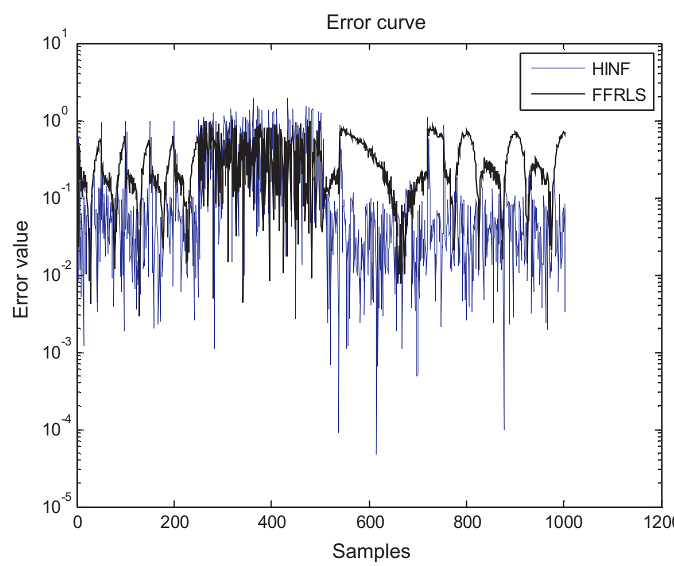

Fig. 12. Absolute error between true and estimated output for input excitation-1 in LLWNN (Example-2).

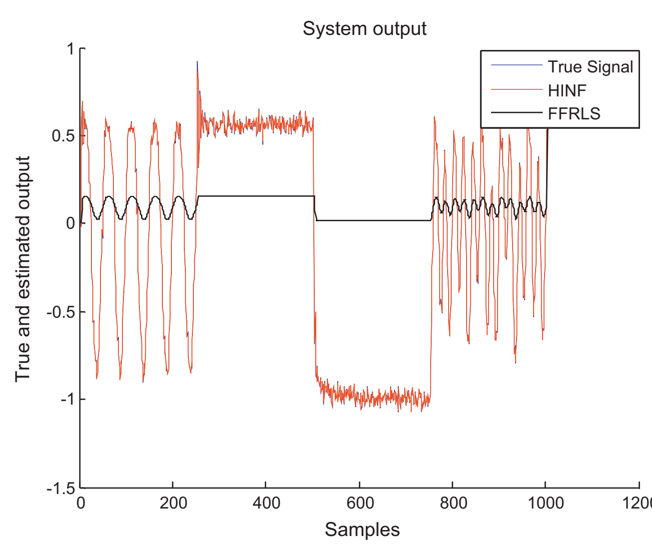

Fig. 13. True and estimated output for input excitation-2 using trigonometric NN (Example-2).

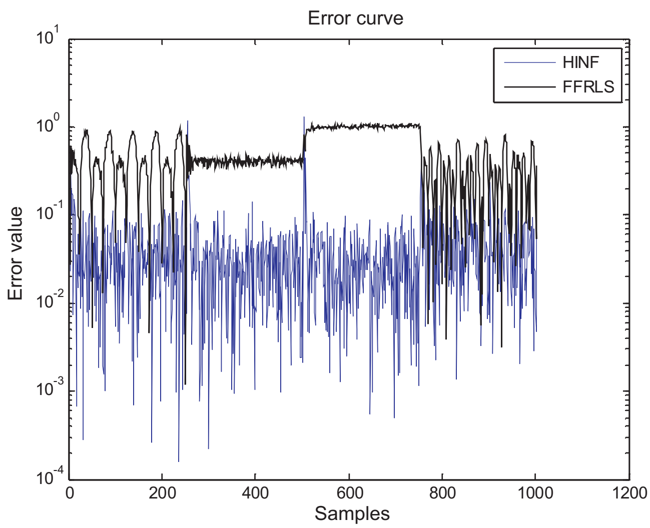

Fig. 14. Absolute error between true and estimated output for input excitation-2 using trigonometric NN (Example-2).

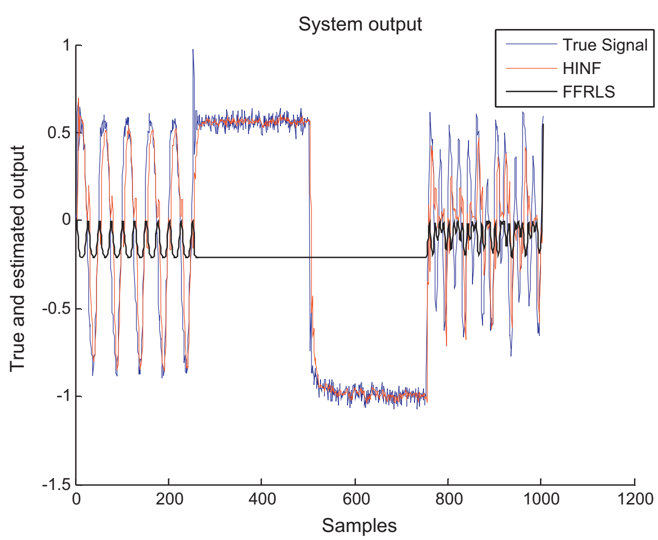

Fig. 15. True and estimated output for input excitation-2 using LLWNN (Example-2).

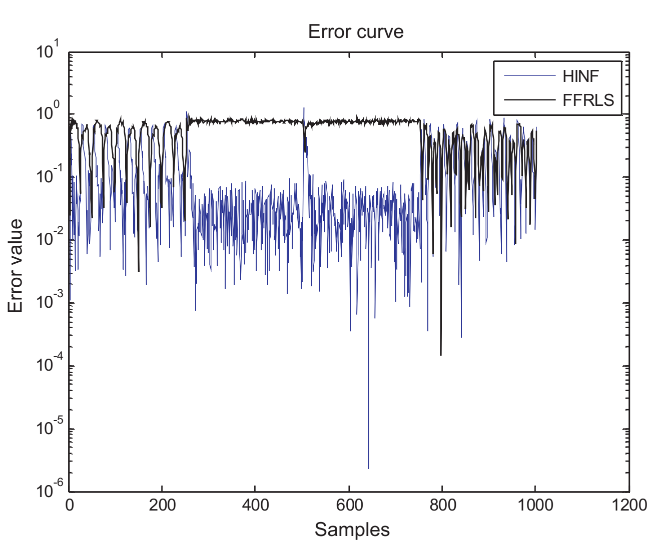

Fig. 16. Absolute error between true and estimated output for input excitation-2 using LLWNN (Example-2).

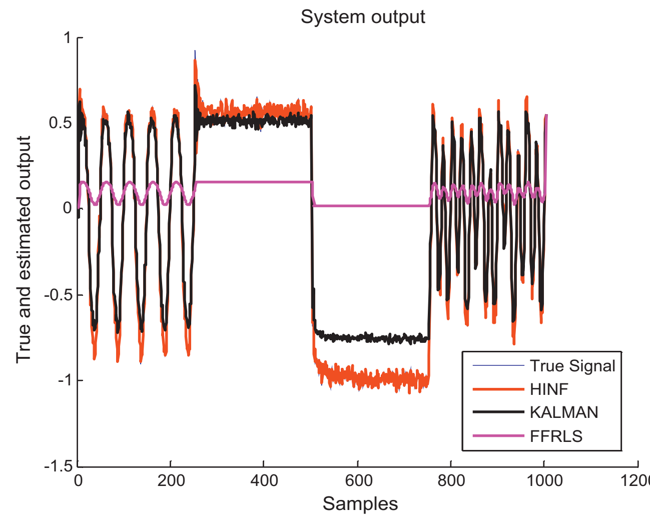

Fig. 17. True and estimated output for input excitation-1 using Legendre polynomial NN (Example-2).

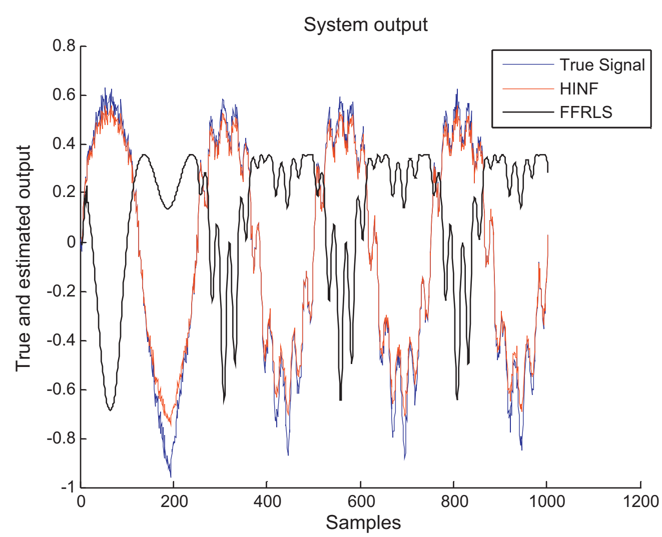

Fig. 18. True and estimated output of SISO plant using trigonometric expansion based NN (Example-3).

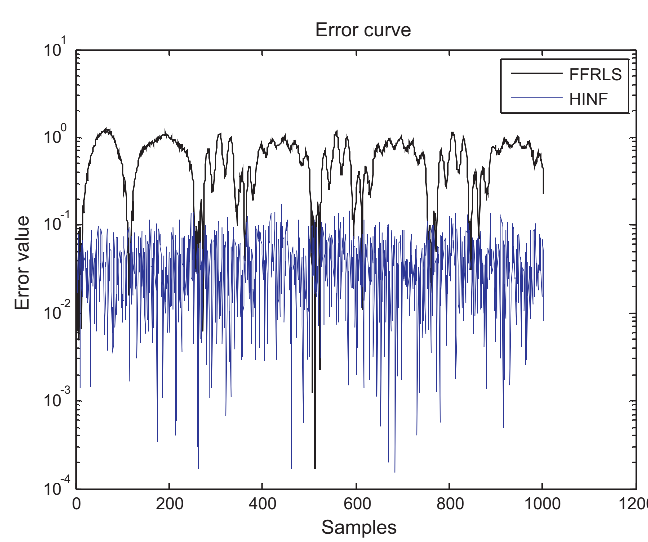

Fig. 19. Absolute error between true and estimated output of SISO plant using trigonometric NN (Example-3).

Table 1 (Chebyshev Neural Network)

No. of Inputs	Inputs chosen	Mean Square Error (MSE)	
		H_∞ filter	FFRLS
2	$y(k-1), u(k-1)$	0.0077	1.0034
3	$y(k-1), y(k-2), u(k-1), u(k-2)$	0.0100	0.8273
6	$y(k-1), y(k-2), y(k-3), u(k-1), u(k-2), u(k-3)$	0.0064	0.4230

Table 2 (Legendre polynomial Based ANN)

No. of Inputs	Inputs chosen	Mean Square Error (MSE)	
		H_∞ filter	FFRLS
2	$y(k-1), u(k-1)$	0.0082	1.1271
3	$y(k-1), y(k-2), u(k-1), u(k-2)$	0.0122	0.3328
6	$y(k-1), y(k-2), y(k-3), u(k-1), u(k-2), u(k-3)$	0.0030	0.3914

Table 3 (Trigonometric Expansion Based ANN)

No. of Inputs	Inputs chosen	Mean Square Error (MSE)	
		H_∞ filter	FFRLS
2	$y(k-1), u(k-1)$	0.0112	1.2025
3	$y(k-1), y(k-2), u(k-1), u(k-2)$	0.0141	0.8287
6	$y(k-1), y(k-2), y(k-3), u(k-1), u(k-2), u(k-3)$	0.0150	0.1654

Table 4 (Local Linear Wavelet NN)

No. of Inputs	Inputs chosen	Mean Square Error (MSE)	
		H_∞ filter	FFRLS
2	$y(k-1), u(k-1)$	0.1125	1.4125
3	$y(k-1), y(k-2), u(k-1), u(k-2)$	0.0146	0.4047
6	$y(k-1), y(k-2), y(k-3), u(k-1), u(k-2), u(k-3)$	0.0173	0.4032

Fig. 20. MSE comparison between H_∞ filter and FFRLS algorithms using different ANN structure.

Square Error) is taken as the performance index of estimation which is calculated as:

$$\text{MSE} = \frac{1}{N}\sum_{k=1}^{N}[y(k) - \hat{y}(k)]^2 \tag{45}$$

The performance of the H_∞ filter is established through comparison of MSE (Mean Square Error) with FFRLS using functional expansions and is presented in separate tabular forms as shown in Fig. 20.

7.1. Example-1

The challenging system to be identified that involves by the following nonlinear and time-varying system:

$$u(k) = au(k-1)^2 + b\frac{u(k-1)}{1+u(k-1)^2} + c\cos(1.2kT) \tag{46}$$

$$y(k) = du(k)^2 \tag{47}$$

where $a = 0.02$, $b = 0.08$, $c = 0.03$, $d = 0.05$, $T = 0.05$.

Input excitation $u(k)$ is generated through rand(.) function. The past input and output samples are expanded using trigonometric functions like sine and cosine functions. The estimated output signal absolute error plots are shown in Figs. 4 and 5.

7.2. Example-2

In this example a second order nonlinear time-varying plant is considered. The process can be described by the following difference equation:

$$y(k) = \frac{y(k-1)y(k-2)y(k-3)u(k-1)(y(k-3)-b(k)) + c(k)u(k)}{a(k) + y(k-2)^2 + y(k-3)^2} \tag{48}$$

where time varying coefficients are given by:

$$a(k) = 1.2 - 0.2\cos\left(\frac{2\pi k}{T}\right); b(k) = 1 - 0.4\sin\left(\frac{2\pi k}{T}\right); c(k) = 1 + 0.4\sin\left(\frac{2\pi k}{T}\right)$$

For example-2 input excitations 1 and 2 are considered for identification.

Input Excitation 1:

$$u(k) = \begin{cases} -0.7 + \mathrm{mod}(k, 50)/40, & k < 250 \\ rands(1, 1), & 250 \le k < 500 \\ 0.7 - \mathrm{mod}(k, 180)/180, & 500 \le k < 750 \\ 0.6\sin\left(\frac{\pi k}{50}\right), & 750 \le k < 1000 \end{cases}$$

Input Excitation 2:

$$u(k) = \begin{cases} \sin(\pi k/25), & k < 250 \\ 1.0, & 250 \le k < 500 \\ -1.0, & 500 \le k < 750 \\ 0.3\sin\left(\frac{\pi k}{25}\right) + 0.1\sin\left(\frac{\pi k}{32}\right) + 0.6\sin\left(\frac{\pi k}{10}\right), & 750 \le k < 1000 \end{cases}$$

Here $y(k-1)$, $y(k-2)$, $y(k-3)$ are one, two and three step delayed outputs of the plant, $u(k)$ and $u(k-1)$ are the current and one step delayed input samples of the plant. Here T is the time span of the test which is chosen as 0.05 for simulation purpose. Estimated outputs and absolute error plots are shown for considering all networks from Figs. 6–17. In Figs. 11 and 17, the estimated plots of extended Kalman filter (EKF) are also included to establish the efficiency of the proposed learning method over other adaptive filtering methods.

7.3. Example-3

SISO plant: In this test model, a single input single output discrete time plant described by:

$$y(k) = f[y(k-1), y(k-2), y(k-3), u(k), u(k-1)] \tag{49}$$

$$\hat{y}(k) = \hat{f}[y(k-1), y(k-2), y(k-3), u(k), u(k-1)] \tag{50}$$

The input excitation is given by:

$$u(k) = \sin(2\pi k/250), \quad 0 < k \le 250$$
$$0.8\sin(2\pi k/250) + 0.2\sin(2\pi k/25), \quad 250 < k \le 1000$$

where the unknown nonlinear function is given by:

$$f[a_1, a_2, a_3, a_4, a_5] = \frac{a_1 a_2 a_3 a_5 (a_3 - 1) + a_4}{1 + a_2^2 + a_3^2}$$

In this example, the input patterns $y(k-1)$, $y(k-2)$, $y(k-3)$, ..., $u(k)$, $u(k-1)$ ... are expanded using trigonometric sine and cosine functions as given in Eq. (11). Same input excitation is given to the actual system and the neural network model. The estimated output and error are given in Figs. 18 and 19.

8. Conclusion

The paper presents an efficient nonlinear system identification scheme using NARX model based neural network structures which are simple and computationally faster. H_∞ filtering is proposed as an robust learning approach for the adaptation of neural weights to identify nonlinear dynamic plants. The tracking results along with MSE comparisons presented in this paper clearly demonstrate the superiority of this approach for dynamic nonlinear system identification in a very high noisy condition.

References

[1] S.A. Billings, S.Y. Fakhouri, Nonlinear system identification using the Hammerstein model, International Journal of Systems Sciences 10 (1979) 567–578.

[2] P. Stoica, T. Soderstorm, Instrumental variable methods for identification of Hammerstein systems, International Journal of Control 35 (1982) 459–476.

[3] F. Giri, F.Z. Chaoui, Y. Rochdi, Parameter identification of a class of Hammerstein plants, Automatica 37 (2001) 749–756.

[4] S.J. Nanada, G. Panda, B. Majhi, Improved identification of Hammerstein plants using new CPSO and IPSO algorithm, Expert Systems with Applications 37 (10) (2010) 6818–6831.

[5] M. Feder, E. Weinstein, Parameter estimation of superimposed signals using the EM algorithm, IEEE Transactions on Acoustics, Speech and Signal Processing 36 (1988) 477–489.

[6] L. Ljung, System Sciences Series, System Identification Theory for the User, second ed., Prentice Hall, Upper Saddle River, NJ, USA, 1999.

[7] L. Ljung, Perspectives on system identification, in: Plenary Talk at the 17th IFAC World Congress, Seoul, Korea, July 6–11, 2008.

[8] K. Narendra, K. Parthasarathy, Identification and control of dynamical systems using neural networks, IEEE Transactions on Neural Networks 1 (1) (1990) 4–27.

[9] S. Chen, S.A. Billings, P.M. Grant, Recursive hybrid algorithm for nonlinear system identification using radial basis function networks, International Journal of Control 55 (5) (1992) 1051–1070.

[10] S.V.T. Elanayar, Y.C. Shin, Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems, IEEE Transactions on Neural Networks 5 (1994) 594–603.

[11] A. Swami, G. Giannakis, S. Shamsunder, Multichannel ARMA processes, IEEE Transaction on Signal Processing 42 (1994) 898–913.

[12] H.-F. Chen, New approach to recursive identification of ARMAX systems, IEEE Transaction on Automatic Control, Mathematical and Computer Modeling 55 (2010) 868–879.

[13] B. Bao, Y. Xu, J. Sheng, R. Ding, Least squares based iterative parameter estimation algorithm for multivariable controlled ARMA system modeling with finite measurement data 53 (2011) 1664–1669.

[14] G. Beligiannis, S. Likothanassis, L. Skarlas, Evolutionary multi-model estimators for ARMA system modeling and time series prediction, Lecture Notes in Computer Science (including subseries lecture notes in AI and Bioinformatics) 2687 (2003) 409–416.

[15] G. Beligiannis, S. Likothanassis, L. Skarlas, A generic applied evolutionary hybrid technique, IEEE Signal Processing Magazine 21 (3) (2004) 28–38.

[16] G. Beligiannis, Evolutionary lainiotis algorithms for system identification: a survey, in: IWSSIP 2005-Proceedings of 12th International Workshop on Systems, Signals and Image Processing, Chalkida, Greece, 22nd–24th September, 2005, pp. 63–67.

[17] Y. Liu, H.Q. Wang, J. Yu, P. Li, Selective recursive kernel learning for online identification of nonlinear systems with NARX form, Journal of Process Control 20 (2) (2010) 181–194.

[18] R.H. Abiyev, O. Kaynak, T. Alshanableh, F. Mamedov, A type-2 neuro-fuzzy system based on clustering and gradient techniques applied to system identification and channel equalization, Applied Soft Computing 11 (2011) 1396–1406.

[19] J.C. Patra, R.N. Pal, B.N. Chatterji, G. Panda, Identification of nonlinear dynamic systems sing functional link artificial neural networks, IEEE Transactions on Systems Man and Cybernetics 29 (2) (1999) 254–262.

[20] P.K. Das, H.P. Satpathy, A.C. Liew, S. Rahman, A real-time short-term load forecasting system using functional link network, IEEE Transactions on Power Systems 2 (2) (1997) 675–680.

[21] J.C. Patra, A.C. Kot, Nonlinear dynamic system identification using Chebyshev functional link artificial neural networks, IEEE Transactions on Systems Man and Cybernetics 32 (4) (2002) 505–511.

[22] S. Purwar, I.N. Kar, A.N. Jha, On-line system identification of complex systems using Chebyshev neural networks, Applied Soft Computing 7 (2007) 364–372.

[23] J.C. Patra, W.C. Chin, P.K. Meher, Legendre-FLANN-based nonlinear equalization in wireless communication system, IEEE International Conference on Systems, Man and Cybernetics (2008) 1826–1831.

[24] S. Chakravarty, P.K. Dash, Dynamic filter weights neural network model integrated with differential evolution for day-ahead price forecasting in energy market, Expert Systems with Applications 38 (2011) 10974–10982.

[25] K. Nishiyama, Robust estimation of a single complex sinusoid in white noise-H_∞ filtering approach, IEEE Transactions on Signal Processing 47 (1999) 2853–2856.

[26] W. Li, Y. Jia, H-infinity filtering for a class of nonlinear discrete-time systems based on unscented transform, Signal Processing 90 (2010) 3301–3307.

[27] H.K. Sahoo, P.K. Dash, N.P. Rath, Frequency estimation of distorted non-stationary signals using complex H_∞ filter, International Journal of Electronics and Communications (AEU) 66 (2012) 267–274.

[28] G.B. Giannakis, E. Serpedin, A bibliography on nonlinear system identification, Signal Processing 81 (3) (2001) 533–580.

[29] I. Petrovic, M. Baotic, N. Peric, Model structure selection for nonlinear system identification using feed forward neural networks, in: IJCNN 2000, Proceedings of the IEEE-INNS-ENNS, vol. 1, 2000, pp. 53–57.
[30] B. Majhi, G. Panda, Development of efficient identification scheme for nonlinear dynamic systems using swarm intelligence techniques, Expert Systems with Applications 37 (2010) 556–566.
[31] Y. Liu, D. Wang, F. Ding, Least square based iterative algorithms for identifying Box–Jenkins models with finite measurement data, Digital Signal Processing 20 (2010) 1458–1467.
[32] K. Reif, S. Günther, E. Yaz, R. Unbehauen, Stochastic stability of the discrete-time extended Kalman filter, IEEE Transaction on Automatic Control 44 (4) (1999) 714–728.