

Recurrent Interval Type-2 Fuzzy Wavelet Neural Network with Stable Learning Algorithm: Application to Model-based Predictive Control

Bibi Elham Fallah Tafti · Mohammad Teshnehlab · Mojtaba Ahmadih Khanesar

Received: date / Accepted: date

Abstract Fuzzy neural networks, with suitable Learning strategy, have been demonstrated as an effective tool for online data modelling. However, it is a challenging task to construct a model to ensure its quality and stability for non-stationary dynamic systems with some uncertainties. To solve this problem, this paper presents a novel identification model based on recurrent interval type-2 fuzzy wavelet neural network (RIT2FWNN) with new learning algorithm. The model benefits from both advantages of recurrent and wavelet neural networks such as use of temporal data and fast convergence properties. The proposed antecedent and consequent parameters update rules are derived using sliding-mode-control-theory. To evaluate the proposed fuzzy model, it is utilized to design a nonlinear model-based predictive controller and is applied for the synchronization of fractional-order time-delay chaotic systems. Using Lyapunov stability analysis, it is shown that all update rules of the parameters are uniformly ultimately bounded. The adaptation laws obtained in this method are very simple and have closed forms. Some stability conditions are derived to prove learning dynamics and asymptotic stability of the network by using an appropriate Lyapunov function. The efficacy

and performance of the proposed method is verified by simulation examples.

Keywords Identification · Sliding mode · Model predictive controller · Fractional order chaotic systems · Synchronization

1 Introduction

In the study of nonlinear dynamical system identification, conventional modeling approaches may not be suitable to be used due to the lack of precise, formal knowledge about the system, strongly nonlinear behavior, a high degree of uncertainty and time-varying characteristics. Most of problems require methods that can handle qualitative and quantitative information with varying precision and complexity. These considerations impose extra demands on the effectiveness of process modeling techniques [31]. Model-free methods like fuzzy logic and neural network are universal estimators that can be applied to approximate the behavior of the nonlinear dynamical systems [21]. Fusion of these techniques into a system enhances the capability of them like fuzzy neural network [8, 18].

Since the output of dynamic systems is a function of time-delayed input and time-delayed output, recurrent neural network (RNN) is suitable choice for identifying their behaviour. These networks have recurrent loops in their topologies that makes it possible to temporarily store information, capture dynamic response of the system and enhance the approximation accuracy of the network [19, 15]. In addition, combination of RNN and type-2 fuzzy sets are applied to various practical applications [11, 19, 17] to enhance noise tolerance of the system; different methods for their parameter tuning are discussed. The recurrent fuzzy structure

Bibi Elham Fallah Tafti
Department of Computer Science, Science and Research
Branch, Islamic Azad university, Tehran, Iran
E-mail: elham.fallah@srbiau.ac.ir

Mohammad Teshnehlab
Faculty of Electrical and Computer Engineer, K. N. Toosi
University of Tech., Tehran, Iran
E-mail: teshnehlab@eetd.kntu.ac.ir

Mojtaba Ahmadih Khanesar
Faculty of Engineering University of Nottingham, Nottingham,
NG7 2RD, U.K
E-mail: ezzma5@exmail.nottingham.ac.uk

that is used in [17] formed an external loop and internal feedback by feeding the rule firing strength of each rule to other rules and itself. The structure benefits from variable-dimensional Kalman filter and gradient descent (GD) algorithm for tuning its parameters. The recurrent Takagi-Sugeno-Kang (TSK) fuzzy structure in [10] obtains from locally feeding the firing strength of each fuzzy rule back to itself. An iterative linear support vector regression algorithm is applied to tune all parameters. In addition interval type-2 fuzzy system applied to represent the complex nonlinear plant in practical industrial systems [26,27]. For example on fault detection problem, the author in [26] utilized it to design an event-triggered fault detection filter for non-linear networked systems which the parameter uncertainty is captured by type-2 fuzzy membership functions.

Due to the number of possible nonlinear interactions in nonlinear dynamical system is theoretically infinite, a nonlinear function should be chosen which is rich enough to describe a nonlinear process with good accuracy. Moreover, certain types of functions can efficiently approximate only certain nonlinear relationships [31]. Some basis functions have important general properties to deal with nonlinearity and uncertainties like wavelet representation.

Wavelet function as the activation function can enhance the advantages of neural networks for faster learning ability and wavelet decomposition for identification purposes. In literature, wavelet neural network (WNN) and its integration with fuzzy logic to determine an optimal definition of premise and consequent part of fuzzy rules, are applied in identification and control of nonlinear dynamical systems [21,5,20]. Discrete wavelet transform and WNN which is trained by back propagation and GD algorithms has also been used to improve the pattern recognition effects of sEMG signals [6].

Stability and convergence are fundamental in numerical analysis for on-line modeling, identification and control tasks. One of the foremost stability analysis techniques is the direct implementation of Lyapunov's stability theory. Another way of designing a robust and stable system is to use the variable structure systems (VSS) theory, which construct the parameter adaptation mechanism and a rigorous stability analysis [30,9].

Sliding mode control (SMC) approach as a class of VSS has high performance in dealing with uncertainties and imprecision. Since robustness and invariance properties to matched uncertainties are the most significant properties of an SMC, the use of it in artificial neural networks (ANN) or fuzzy neural networks (FNN) can ensure convergence and stability of the learning algorithm [33]. As the result has shown in [30], SMC

improves the performance of systems based on soft-computing techniques which utilize the gradient-based training strategies.

In literature, several research works use the diffusion of SMC approach into ANNs. In [12] a sliding mode incremental learning algorithm is used for tuning the parameters of interval type-2 FNN (IT2FNN) where an adaptive learning rate with an adaptation law is derived. An adaptive controller for speed control of induction motor which utilizes IT2FNN and SMC-based learning algorithm is proposed in [22]. In another study, a new learning algorithm for radial basis function neural networks is demonstrated that is based on applying fast terminal sliding mode [13]. In [34], a novel identification and control scheme using multi-time scale recurrent high-order neural networks is presented. The scheme uses modified optimal bounded ellipsoid based weight's updating laws to identify a nonlinear systems.

Recently, the synchronization of the fractional order chaotic systems have been studied for many applications in science and technology. It has been observed that many applications arising in various fields of science and engineering are described more accurately by fractional order differential equations. Fractional calculus is a mathematical topic which deals with derivatives and integrations with non-integer order and applied to model non-linear biological systems with complex behavior and long-term memory [2].

Chaos is a nonlinear and deterministic phenomenon which characterized by a subset of features comprising having an unusual sensitivity to initial states, not being periodic, having fractal structures, and being governed by one or more control parameters [7]. On the other hands, the chaotic dynamics of fractional-order differential systems has gained attraction in the investigation of dynamical systems, such as the fractional order of Chua's circuit, Duffing system [1], Lu system, Chen system [28].

In [4], a direct adaptive fuzzy controller is designed to obtain a generalized projective synchronization of two different incommensurate fractional-order chaotic systems in the presence of both uncertain dynamics and external disturbances. Lin and Lee developed an adaptive fuzzy-control scheme incorporating with SMC approach to synchronize two nonlinear fractional-order Duffing-Holmes chaotic systems [16]. They investigated the effect of delay on the chaotic behavior of the fractional-order system for the first time. Mohammadzadeh et.al formulated a robust nonlinear model-based predictive control for synchronization of the fractional-order chaotic systems [24]. By considering a fractional-order Lyapunov function, Mohadeszadeh and Delavari designed an adaptive finite-time sliding mode control for chaos synchro-

nization between two identical and non-identical fractional-order hyper-chaotic systems [23]. The robust adaptive interval type-2 fuzzy control strategy incorporating Lyapunov stability criterion and H_∞ synchronization performance is studied in [14].

The models of real-world systems is not known completely and/or their accuracy are affected by nonlinear and time-varying behaviour which can be originated from actual high degrees of uncertainties about the plant or from the plant dynamics, external disturbances and time-varying parameters. Identification performance of such applications is influenced by these factors. This study presents a new robust nonlinear and stable identification model RIT2FWNN to address these issues. In order to show the capabilities of the proposed model in such conditions, it is used to identify a second-order nonlinear time-varying plant and as a model in MPC to estimate the tracking error online. The controller is applied for the synchronization of uncertain fractional order chaotic systems and the robustness of its synchronization in the presence of external disturbances, and approximation errors are investigated. The model performance in handling the uncertainties and imprecision can be achieved by utilization of recurrent IT2FNN and wavelet activation function. For the parameter adaptation of the RIT2FWNN, sliding mode theory-based supervised online training algorithm is elaborated. This learning algorithm is shown to be stable in the sense of Lyapunov stability theory which derives mathematical stability proofs to ensure finite time stability in RIT2FWNN. Moreover, fast convergence learning speed and the robustness of recurrent proposed structure is guaranteed. The main contributions of this manuscript are as follows:

- The proposed structure of RIT2FWNN includes wavelet activation function with its parameters being tuned by a novel sliding-mode-control-theory learning algorithm.
- Utilizing the Lyapunov stability theorem which is applied to obtain the parameter tuning algorithms. As it is expected, it is shown that the error value will be smaller in cases where the number of rules and input variables are more.
- Since the performance of MPC highly depends on the accuracy of the model, a robust model-based predictive controller scheme is developed with the aid of the proposed model.
- The robustness of the developed model-based predictive controller is examined in the case there exist uncertain dynamics and time delays in the dynamic system.

Different symbols defined in this paper are listed in Table 1.

The organization of this paper is as follows. In Section 2, problem statement and preliminaries are given. The structure of RIT2FWNN and the novel SMC-theory-based learning rules are introduced in Section 3. The enhanced model-based predictive controller is presented in Section 4. To demonstrate the validation of the proposed method, synchronization of two fractional-order time-delay chaotic systems and identification of a nonlinear dynamic system are investigated in Section 5. Finally, the conclusions are given in Section 6.

Table 1: Symbols defined in this paper

Symbol	Definition	Symbol	Definition
D_t^α	α -order Caputo differential operator	c	Center of MF
$\Gamma(\cdot)$	Euler's gamma function	$\bar{\sigma}$	Upper standard deviation of MF
$f(X, t)$	Bounded smooth nonlinear function	$\underline{\sigma}$	Lower standard deviation of MF
$\omega(t)$	External bounded disturbance	ξ	Recurrent parameter
y_d	Reference signal	$\underline{\theta}$	Lower weights of recurrent node
e	Tracking error	$\bar{\theta}$	Upper weights of recurrent node
y	System output	\underline{w}	Lower firing strength
x	System's states	\bar{w}	Upper firing strength
u_s	Predictive control signal	$\frac{\underline{w}}{\bar{w}}$	Normalized values of \underline{w}
I	Number of input variables	$\frac{\bar{w}}{\bar{w}}$	Normalized values of \bar{w}
N	Number of fuzzy rules	q	Design parameter
h	Number of MF	S	Sliding surface
f_r	Output of r th fuzzy rule	α	Adaptive learning rate
\bar{A}	Type-2 fuzzy MF	ν	Small positive number
ρ	Weight coefficient	γ	Learning rate for (α)
a	Dilation	R, Q	Weighting matrices
b	Translation	N_c	Control horizon
$\psi(x)$	Mother wavelet function	N_p	Prediction horizon
$\bar{\mu}$	Upper membership degree	ϑ	Step response coefficient
$\underline{\mu}$	Lower membership degree	J	Cost function
e_r	Free response	e_o	Forced response

2 Problem Statement and Preliminaries For Fractional-Order Systems

The commonly used definitions in literature for Fractional-Order operator are Grunwald-Letnikov, Riemann-Liouville, and Caputo definitions. The last one is introduced for engineering applications because its Laplace transform requires integer-order derivatives for the initial conditions. The Caputo's fractional derivative of a function $x(t)$ with respect to time is defined as follows [29]:

$$D_t^\alpha = \frac{1}{\Gamma(m-\alpha)} \int_0^t (t-\tau)^{-\alpha+m-1} x^{(m)}(\tau) d\tau \quad (1)$$

where $m = [\alpha] + 1$, $[\alpha]$ is the integer part of α , D_t^α is called the α -order Caputo differential operator, and $\Gamma(\cdot)$ is the well-known Euler's gamma function:

$$\Gamma(P) = \int_0^\infty t^{P-1} e^{-t} dt; \quad \text{with } \Gamma(P+1) = P\Gamma(P) \quad (2)$$

This function can be considered as an extension of the factorial to real number arguments. The mathematical model of the fractional-order chaotic nonlinear systems to be discussed in this paper can be described as:

$$\begin{cases} D_t^{(\alpha_1)} x_1 = x_2 \\ \vdots \\ D_t^{(\alpha_{n-1})} x_{n-1} = x_n \\ D_t^{(\alpha_n)} x_n = f(X, t) + u + \omega(t) \\ y = x_1 \end{cases} \quad (3)$$

where $X = [x_1, x_2, \dots, x_n]^T = [x, x^{(\alpha)}, x^{(2\alpha)}, \dots, x^{(n-1)\alpha}]$ is the system's states vector, $y \in R$ outputs, $u \in R$ is the control input with the initial conditions, $u(0) = 0$ and $y(0) = 0$. If $\alpha_1 = \alpha_2 = \dots = \alpha_n = \alpha$ the above system is called a commensurate order system. Then equivalent form of the above system is specified as:

$$\begin{aligned} x^{n\alpha} &= f(X, t) + u + \omega(t) \\ y &= x_1 \end{aligned} \quad (4)$$

The unknown function $f(X, t)$ is a bounded smooth nonlinear function which specifies system dynamics and $\omega(t)$ is the external bounded disturbance. The main objective is to force the system output y to follow a given bounded reference signal y_d while assuring under certain constraint, all involved signals are bounded. To quantify this objective, the reference signal vector y_d and the tracking error vector e are defined as,

$$y_d = [y_d, y_d^{(\alpha)}, \dots, y_d^{(n-1)\alpha}]^T \quad (5)$$

$$e = y_d - y = [e, e^{(\alpha)}, \dots, e^{(n-1)\alpha}]^T \quad (6)$$

$$e^{(i\alpha)} = y_d^{(i\alpha)} - y^{(i\alpha)}$$

By substituting (6) into (4), the control system in the state space domain is obtained as follows:

$$\begin{aligned} x^{n\alpha} &= Ax + B[f(X) + u] \\ y &= C^T x \end{aligned} \quad (7)$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \ddots & \dots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ -k_1 & -k_2 & -k_3 & -k_4 & \dots & -k_n \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}. \quad (8)$$

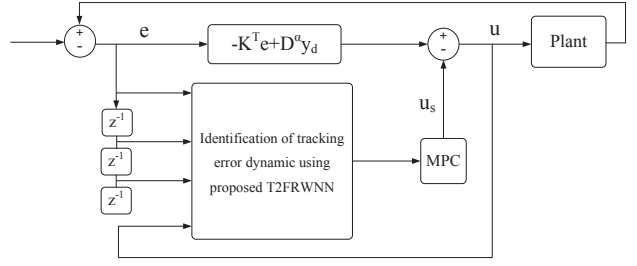


Fig. 1: The designed control block diagram

The equivalent control effort can be considered as follows:

$$u = -K^T e + D_t^\alpha y_d + u_s \quad (9)$$

where u_s is predictive control signal in previous sample times which is designed based on model predictive control. $K = (k_1, k_2, \dots, k_n)^T \in R$ is chosen such that the stability condition $|\arg(\text{eig}(A))| > q\pi/2$, $0 < q < 1$ is satisfied, $\text{eig}(A)$ is the eigenvalues of the system state matrix given in (7).

To design u_s , dynamic of the tracking error is modeled by the proposed RIT2FWNN with its premise and consequent part as well adaptive learning rate are updated using the proposed online learning algorithm based on SMC. Fig.1 shows a conceptual diagram of the control scheme.

3 Proposed Recurrent Interval Type-2 Fuzzy Wavelet Neural Network structure

The structure of RIT2FWNN implements a recurrent wavelet fuzzy model is organized into layers whose consequent part is nonlinear function of the input variables. Each rule in RIT2FWNN has the following form:

R_r : If x_1 is $\tilde{A}_{1j} \dots x_i$ is \tilde{A}_{ih} and \dots and x_I is \tilde{A}_{II} then:

$$f_r = \sum_{i=1}^I \rho_{ri} |a_{ri}|^{-\frac{1}{2}} (1 - z_{ri}^2) e^{-\frac{z_{ri}^2}{2}} \quad (10)$$

where I is the number of input variables, h is the number of membership function and f_r is the output of r th rule ($r = \{1, 2, \dots, N\}$). \tilde{A}_{ik} is the k th type-2 fuzzy membership function (MF) related to i th input variable. ρ is the weight coefficient between the input and the hidden layer. Consequent part of above rule involved a wavelet function of input variables. Wavelets are defined by a family of functions a and b ($a > 0, b \in R$) as follows:

$$\Psi_r(x) = |a_r|^{-\frac{1}{2}} \psi\left(\frac{X - b_r}{a_r}\right), \quad a_r \neq 0 \quad (11)$$

where $\Psi_r(x)$ demonstrates the family of wavelets obtained from the single $\psi(x)$ function, called as a mother wavelet and localized in both time and frequency space, by dilation and translation, $a_r = \{a_{r1}, a_{r2}, \dots, a_{rI}\}$ and $b_r = \{b_{r1}, b_{r2}, \dots, b_{rI}\}$, respectively and $x_r = \{x_1, x_2, \dots, x_I\}$ are input variables. In this manuscript, among several families of wavelets, Mexican Hat is considered as mother wavelet function. It is derived from a function that is proportional to the second derivative function of the Gaussian probability density function.

$$\Psi_r(z) = \sum_{i=1}^I |a_{ri}|^{-\frac{1}{2}} (1 - z_{ri}^2) e^{-\frac{z_{ri}^2}{2}} \quad (12)$$

where, $z_{ri} = \frac{x_i - b_{ri}}{a_{ri}}$.

RIT2FWNN consists of seven layers. In the first layer, nodes representing input linguistic variables are fed into the network. In the second layer, each node corresponds to one linguistic term. For each entering input variable, type-2 MFs are used which have uncertain standard deviation and fixed center. The membership degree $\bar{\mu}_{ik}(x_i)$ and $\underline{\mu}_{ik}(x_i)$ are calculated according to (13):

$$\begin{aligned} \bar{\mu}_{ik}(x_i) &= \exp\left(-\frac{1}{2} \frac{(x_i + \xi_{ik} - c_{ik})^2}{\bar{\sigma}_{ik}^2}\right) \\ \underline{\mu}_{ik}(x_i) &= \exp\left(-\frac{1}{2} \frac{(x_i + \xi_{ik} - c_{ik})^2}{\underline{\sigma}_{ik}^2}\right) \end{aligned} \quad (13)$$

where c_{ik} is center of type-2 MF, $\bar{\sigma}_{ik}$ and $\underline{\sigma}_{ik}$ are the upper and lower standard deviation of the k th type-2 MF of i th input. Moreover, ξ_{ik} is the recurrent parameter defined as (14) which store the past information of the network. It is to be noted that the feedback weights of nodes in layer 2 are interval values.

$$\xi_{ik} = \frac{\theta_{ik} \underline{\mu}_{ik}(t-1) + \bar{\theta}_{ik} \bar{\mu}_{ik}(t-1)}{\theta_{ik} + \bar{\theta}_{ik}} \quad (14)$$

where θ_{ik} and $\bar{\theta}_{ik}$ are considered as feedback weights of the nodes. Nodes in the third layer correspond to one fuzzy rule and perform a fuzzy meet operation on inputs from layer 2 to obtain upper and lower firing strength as follows:

$$\begin{aligned} \bar{w}_r &= \bar{\mu}_{\bar{A}1}(x_1) * \bar{\mu}_{\bar{A}2}(x_2) * \dots * \bar{\mu}_{\bar{A}I}(x_I) \\ \underline{w}_r &= \underline{\mu}_{\bar{A}1}(x_1) * \underline{\mu}_{\bar{A}2}(x_2) * \dots * \underline{\mu}_{\bar{A}I}(x_I) \end{aligned} \quad (15)$$

Layer 4 determines the normalized values of the lower and the upper firing strength corresponding to each node in layer 3:

$$\tilde{w}_r = \frac{w_r}{\sum_{r=1}^N w_r} \quad \text{and} \quad \tilde{\bar{w}}_r = \frac{\bar{w}_r}{\sum_{r=1}^N \bar{w}_r} \quad (16)$$

Layer 5 is consequent layer. In this layer, Nodes compute the product of normalized firing strength $\tilde{w}_r, \tilde{\bar{w}}_r$

and wavelet function of input variables. Layer 6 involves two summation blocks, one of them is for upper and the other is for lower outputs of the previous layer. Layer 7 calculates the output of the network using (17):

$$\begin{aligned} y_N &= q \sum_{r=1}^N f_r \tilde{w}_r + (1-q) \sum_{r=1}^N f_r \tilde{\bar{w}}_r = q \sum_{r=1}^N \rho_r |a_r|^{-\frac{1}{2}} \psi\left(\frac{X - b_r}{a_r}\right) \tilde{w}_r \\ &+ (1-q) \sum_{r=1}^N \rho_r |a_r|^{-\frac{1}{2}} \psi\left(\frac{X - b_r}{a_r}\right) \tilde{\bar{w}}_r \end{aligned} \quad (17)$$

where q is the design parameter which enables to adjust the contribution of the lower or the upper values of rules depending on identification requirements for the system [3]. In literature, q has either been considered to be a constant or a time-varying parameter [12]. In this paper, the adaptation laws for the parameters and the proof of the stability of the learning process are given using a time-varying q .

Remark 1: Main improvement made in the proposed approach with respect to existing approaches in literature [26,27] are use of wavelet functions which makes it possible to benefit from a local nonlinear general function approximator. This in turn will improve function approximation property of the system and provides means for a tight control system. On the other hand, the existence of recurrent term in type-2 fuzzy membership functions allows the use of a short term memory which captures the dynamics of the system and improves the performance of controller.

3.1 Sliding Mode on-line Learning Algorithm

Based on the principles of the SMC theory, the zero value of the learning error coordinate can be defined as a time-varying sliding surface, i.e.

$$S(e(t)) = e(t) = y_N(t) - y(t) = 0 \quad (18)$$

which is the condition that guarantees when the system is on the sliding surface S the RIT2FWNN output $y_N(t)$ will track the desired output signal $y(t)$ for all time $t > t_h$ where t_h is the hitting time of $e(t) = 0$.

Definition 1 A sliding motion will have place on a sliding manifold $S(e(t)) = e(t) = 0$, after time t_h if the condition $S(t)\dot{S}(t) < 0$ is true for all t in some nontrivial semi open subinterval of time of the form $[t, t_h) \subset (0, t_h)$.

The online learning algorithm for the adaptation of the parameters of RIT2FWNN should be derived in such a way that the sliding mode condition of the above definition will be enforced.

Theorem 1 *If the learning algorithm for the parameters of upper and lower Gaussian MFs and weights of feedback loops are chosen as follows:*

$$\dot{c}_{ik} = \dot{x}_i + (x_i - c_{ik} + \xi_{ik})\alpha \operatorname{sgn}(e) \quad (19)$$

$$\dot{\underline{\sigma}}_{ik} = -(\underline{\sigma}_{ik} + \frac{\underline{\sigma}_{ik}^3}{(x_i - c_{ik} + \xi_{ik})^2})\alpha \operatorname{sgn}(e) \quad (20)$$

$$\dot{\bar{\sigma}}_{ik} = -(\bar{\sigma}_{ik} + \frac{\bar{\sigma}_{ik}^3}{(x_i - c_{ik} + \xi_{ik})^2})\alpha \operatorname{sgn}(e) \quad (21)$$

$$\dot{\underline{\theta}}_{ik} = -(x_i - c_{ik} + \xi_{ik})^{-1}(\frac{2}{\bar{\sigma}_{ik} + \underline{\sigma}_{ik}})^{-2} (\frac{\underline{\mu}_{ik}(t-1) - \xi_{ik}}{\underline{\theta}_{ik} + \bar{\theta}_{ik}})^{-1}\alpha \operatorname{sgn}(e) \quad (22)$$

$$\dot{\bar{\theta}}_{ik} = -(x_i - c_{ik} + \xi_{ik})^{-1}(\frac{2}{\bar{\sigma}_{ik} + \underline{\sigma}_{ik}})^{-2} (\frac{\bar{\mu}_{ik}(t-1) - \xi_{ik}}{\underline{\theta}_{ik} + \bar{\theta}_{ik}})^{-1}\alpha \operatorname{sgn}(e) \quad (23)$$

and the adaptation of the parameters of wavelet functions and coefficient in the consequents part of fuzzy rules are chosen as follows:

$$\dot{a}_{ri} = -\frac{1}{\Psi_{ri}(z)\rho_{ri}}(-\frac{1}{2}a_{ri}^{-1} + \frac{z_{ri}^2}{a_{ri}}\frac{3 - z_{ri}^2}{1 - z_{ri}^2})^{-1} \frac{q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r}{(q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r)^T(q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r)}\alpha \operatorname{sgn}(e) \quad (24)$$

$$\dot{b}_{ri} = -\frac{1}{\Psi_{ri}(z)\rho_{ri}}(\frac{z_{ri}}{a_{ri}}\frac{3 - z_{ri}^2}{1 - z_{ri}^2})^{-1} \frac{q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r}{(q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r)^T(q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r)}\alpha \operatorname{sgn}(e) \quad (25)$$

$$\dot{\rho}_{ri} = -\frac{1}{\Psi_{ri}(z)}\frac{q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r}{(q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r)^T(q\tilde{w}_r + (1-q)\tilde{\bar{w}}_r)}\alpha \operatorname{sgn}(e) \quad (26)$$

and the parameter q is updated as follows:

$$\dot{q} = -\frac{1}{F(\tilde{W} - \tilde{\bar{W}})^T}\alpha \operatorname{sgn}(e) \quad (27)$$

and the adaptive learning rate (α) changes as follows:

$$\dot{\alpha} = (2NI + N + 1)\gamma|e| - \nu\gamma\alpha \quad \nu, \gamma > 0 \quad (28)$$

The parameter γ is considered as the learning rate for the adaptive learning rate α which has a small positive real value. The parameter ν should be very small not to interrupt the adaptation mechanism. Then, given an arbitrary initial condition $e(0)$, the learning error $e(t)$ will converge to zero within a finite time t_h .

Proof: The reader is referred to the Appendix. According to the following equation which obtained from the proof, the error value will be smaller in cases where the number of rules and input variables are more.

$$|e| \leq \frac{\alpha^*\nu}{2(3IN + 1)} \quad (29)$$

In equation (28) the parameter γ is the learning rate for the adaptive learning rate α . Since the first term of this equation is positive, the second term is considered to avoid bursting in parameter α . The value of the parameter ν should be adjusted in such a way that the adaptation mechanism not to be disturbed.

Despite aforementioned advantages of simplicity and robustness of SMC, it suffers from a technical problem, called chattering. Sensitivity to noise is another phenomena caused by the presence of signum function in the structure of SMC. In this manuscript, in order to alleviate the problem associated with the chattering and sensitivity to noise, a continuous approximation method is used which smooths the discontinuity caused by the signum function as follows:

$$\operatorname{sign}(e(t)) \approx \frac{e(t)}{|e(t)| + \delta} \quad (30)$$

where δ is a small positive number [33].

4 Model Predictive Control Design

The target of the following equation (31) as the cost function is to solve an enhanced model-based predictive control:

$$\min_{u_s(k), \dots, u_s(k+N_c)} J = \sum_{k=t}^{t+N_p} \hat{e}(k)^T Q \hat{e}(k) + \Delta u_s(k)^T R \Delta u_s(k) \quad (31)$$

$$\hat{e}(t) = \text{RIT2FWNN}[e(t-k_1), \dots, e(t-k_l), u_s(t-1)]$$

where u_s is predictive control signal, $e(t-k_l)$, $l = 1, \dots, r$ are the tracking errors, $\Delta u_s(t) = u_s(t) - u_s(t-1)$, R and Q are positive definite weighting matrices. N_c and N_p are control and prediction horizon. In order to minimize the cost function (31), the output predictions over the horizon must be computed. RIT2FWNN is used to estimate the tracking error $e(t)$. The consequent parameters of RIT2FWNN are tuned based on the updating rule proposed in previous section. By considering the Diophantine equation corresponding to the prediction for $\hat{e}(t+k+1|t)$:

$$\hat{e}(t+k|t) = G_k(z^{-1})\Delta u_s(t+k-1) + G_{kp}(z^{-1})\Delta u_s(t-1) + F_k(z^{-1})e(t) \quad (32)$$

The last two terms of equation (32) depend on past values of the process output and input variables. That is, they correspond to the free response e_r of the process considered if the control signals are kept constant and are computed as follows:

$$e_r(t+k|t) = \text{RIT2FWNN}[e(t+k-1), \dots, e(t+k-r), u_s(t-1)] \quad (33)$$

The first term of equation (32) depends only on future values of the control signal and can be interpreted as the forced response which is obtained by (34). That is, the response obtained when the initial conditions are zero $e(t-k) = 0$, $\Delta u_s(t-k) = 0$.

$$e_o(t+k|t) = \sum_{i=0}^{k-1} \vartheta_i \Delta u_s(t+k-i+1|t) \quad (34)$$

where ϑ_i , $i = 0, \dots, k-1$ are the step response coefficients that are obtained using unit step on RIT2FWNN model. Step response can be estimated as follows [25]:

$$g(t-1) = \frac{\hat{e}_{step}(t+k|t) - \hat{e}_r(t+k|t)}{du_s(t)} \quad (35)$$

where $du_s(t)$ is the step size and $\hat{e}_{step}(t+k|t)$ is calculated by RIT2FWNN like as e_r . Equation (32) can be rewritten as follows:

$$\hat{e}(t+k|t) = G_k(z^{-1})\Delta u(t+k-1) + f_k \quad (36)$$

such that $f_k = G_{kp}(z^{-1})\Delta u_s(t-1) + F_k(z^{-1})e(t)$. Then the step ahead prediction of the system output on data up to time N_p in matrix form described as follows:

$$\begin{bmatrix} \hat{e}(t+1|t) \\ \hat{e}(t+2|t) \\ \vdots \\ \hat{e}(t+N_p|t) \end{bmatrix} = \begin{bmatrix} G_0 & 0 & \cdots & 0 \\ G_1 & G_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ G_{N_p-1} & G_{N_p-2} & \cdots & G_0 \end{bmatrix} \begin{bmatrix} \Delta u(t) \\ \Delta u(t+1) \\ \vdots \\ \Delta u(t+N_p-1) \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_{N_p} \end{bmatrix}$$

The predictions can be expressed in condensed form as follows:

$$E = G\Delta U_s + e_r \quad (37)$$

The optimization problem (31) can be written as follows:

$$\begin{aligned} J &= E^T Q E + \Delta U_s^T R \Delta U_s \\ &= (G\Delta U_s + e_r)^T Q (G\Delta U_s + e_r) + \Delta U_s^T R \Delta U_s \end{aligned} \quad (38)$$

By making the gradient of J in (38) equal to zero:

$$\frac{\partial J}{\partial \Delta U_s} = 2G^T Q f_r + 2(G^T Q G + R)\Delta U_s \quad (39)$$

Then the optimum is:

$$\Delta U_s = (G^T Q G + R)^{-1} G^T Q e_r \quad (40)$$

5 Simulation Studies and Discussions

In this section, two illustrative examples are presented to demonstrate the applicability and feasibility of the proposed method and to confirm the theoretical results. In these examples, the designed controller is applied to synchronize two identical and non-identical uncertain fractional-order chaotic systems in the presence of external disturbances. Besides, to analyze how the proposed structure measures the mathematical model of a system from measurements of the system inputs and outputs, an example of nonlinear system identification is considered.

Example 1: As the first example, the proposed identification model, is applied to identify a second-order nonlinear time-varying plant [12] described as:

$$y(k) = \frac{x_1 x_2 + x_3}{x_4} \quad (41)$$

where $x(1) = y(k-1)y(k-2)y(k-3)u(k-1)$, $x(2) = y(k-3) - b(k)$, $x(3) = c(k)u(k)$ and $x(4) = a(k) + y(k-2)^2 + y(k-3)^2$. The time-varying parameters a , b and c are defined as:

$$\begin{aligned} a(k) &= 1.2 - 0.2\cos(2\pi k/T) \\ b(k) &= 1 - 0.4\sin(2\pi k/T) \\ c(k) &= 1 + 0.4\sin(2\pi k/T) \end{aligned} \quad (42)$$

where $T = 1000$ is the time span of the test. This example have been implemented with three inputs which are fuzzified with three Gaussian type-2 MFs with a fixed center and uncertain standard deviation. The inputs are the delayed signal from the plant output- $y(t-T_0)$, $y(t-2T_0)$ and input signal $u(t)$, with the period of discretization selected as $T_0 = 1$ ms. The incoming network signals have been normalized to be in the range $[-1, 1]$. The input signal $u(k)$ is as follows:

$$u(k) = \begin{cases} \sin(\pi k/25), & k < 250. \\ 1.0, & 250 \leq k < 500. \\ -1.0, & 500 \leq k < 750. \\ 0.3\sin(\pi k/25) + 0.1\sin(\pi k/32) \\ \quad + 0.6\sin(\pi k/10), & 750 \leq k < 1000. \end{cases} \quad (43)$$

Fig. 2(a) compares the plant output with the output yielded by RIT2FWNN. 2(b) shows the root mean square error (RMSE) values versus each epoch number which illustrates the proposed SMC-based learning algorithm is stable. Figs. 2(c) and 2(d) show the adaptation of parameter α and q , respectively. Figs.2(c) shows the learning rate α will be stable at infinity. Table 2 compares the performance of RIT2FWNN with two different approaches and [12]. These results reveal

the benefit of using recurrent structure and SMC-based learning algorithm of RIT2FWNN by achieving smaller RMSE values. It is to be noted that by increasing the nonlinearity of the RIT2FWNN by using wavelet functions in the consequents part of fuzzy rules and feedback loops on type-2 MFs, the accuracy of other learning methods such as gradient-based methods are lowered considerably.

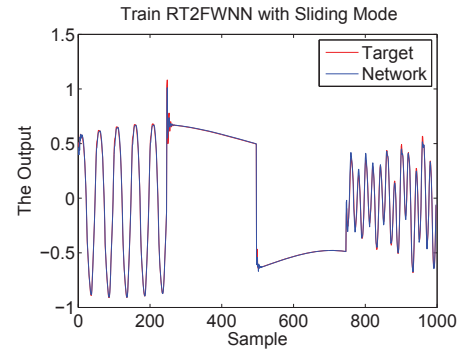
Furthermore, the accuracies of extended Kalman filter and the proposed SMC-based learning method are similar. As previously discussed, the adaptation laws obtained in the proposed SMC-based learning method have explicit forms which do not have any derivative computations of the output of RIT2FWNN concerning trainable parameters and manipulation of some high order matrices. This is the main reason why the computation time for extended Kalman filter and gradient-based methods is higher than SMC-based methods (see Table 2). Additionally, further results are provided to demonstrate the benefits of the proposed model with its type-1 counterpart. As can be seen, the proposed type-2 model achieves smaller test and train RMSE than that by the type-1.

Table 2: The RMSE values of different methods in Ex.1.

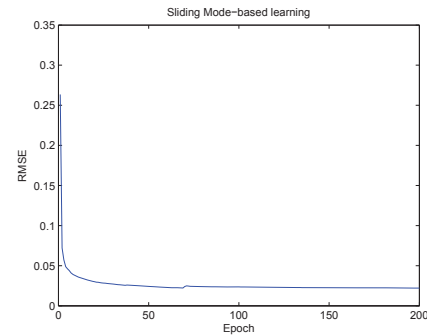
Training methods	GD	EKF	SMC-based	proposed type-2 method	proposed type-1 method
Train RMSE	0.0532	0.0217	0.0342	0.0263	0.0338
Test RMSE	0.0761	0.0245	0.0280	0.0207	0.0251
Time	136.322	312.129	85.3026	135.437	108.371

Remark 2: As the simulation results show, the proposed model have better RMSE results than others. However it is more complicated in network structure. The best advantages of the proposed RIT2FWNN with respect to other such as [12], is the rate of convergence. Most controllers in real world application are based on model and it is needed to be achieved faster. Hence, the degree of freedom in the proposed model is increased to obtain a faster convergence speed. Furthermore, several experiments have been done on the speed of convergence of different models. The average arising times are approximately 29 and 38 seconds for the proposed and [12] respectively, which confirm the favourable convergence behaviour of the proposed model versus time.

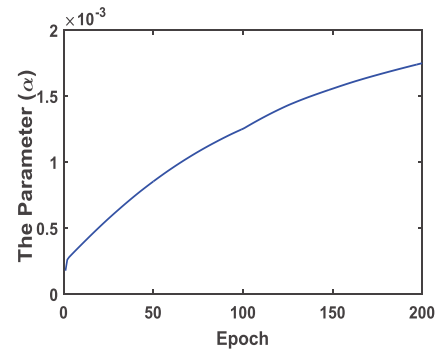
Example 2: As the second example, we will verify the effectiveness of the proposed method to synchronize two different uncertain fractional-order Duffing-Holmes chaotic time-delay systems [1]. The master and slave



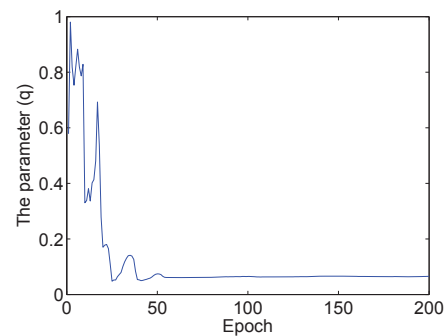
(a)



(b)



(c)



(d)

Fig. 2: (a) Plant output of example 1 and the RIT2FWNN system. (b) RMSE values during the online identification procedure. (c) Adaptation of the learning rate (d) Adaptation of parameter q .

systems dynamics are given as follows, respectively:

$$\begin{cases} D^\alpha x_1(t) = 2.5x_2(t) \\ D^\alpha x_2(t) = -(\frac{1}{2.5}x_1(t))^3 - \frac{1}{2.5}x_1(t) - 0.1x_2(t) \\ \quad + 0.01x_1(t-0.001) + 0.01x_1(t)^2(t-0.001) \\ \quad + 0.01x_2(t-0.001) + 25\cos(1.29t). \end{cases} \quad (44)$$

$$\begin{cases} D^\alpha y_1(t) = 2.2y_2(t) \\ D^\alpha y_2(t) = -(\frac{1}{2.0}y_1(t))^3 - \frac{1}{1.8}y_1(t) - 0.1y_2(t) \\ \quad + 0.01y_1(t-0.001) + 0.01y_1(t)^2(t-0.001) \\ \quad + 0.01y_2(t-0.001) + 25\cos(1.29t) + \omega(t) + u(t). \end{cases} \quad (45)$$

where $\omega(t) = 0.7\sin(t)$ is the external bounded disturbance, and $u(t)$ is control input of the slave system. The initial conditions are $x(0) = [x_1(0), x_2(0)]^T = [0, 0]$ and $y(0) = [y_1(0), y_2(0)]^T = [1, -1]$. The simulation sample time is 0.001 and the fractional derivative order is considered as $\alpha = 0.98$. The control effort of the slave can be obtained as follows:

$$u = -K^T e + D^\alpha x_2 + u_s \quad (46)$$

where $K = [900, 30]$. The following cost function is considered to achieve the control signal u_s :

$$\min_{u_s(k), \dots, u_s(k+10)} J = \sum_{k-t}^{t+10} \hat{e}(t)^2 + \Delta u_s(t)^2 \quad (47)$$

$$\hat{e}(t) = \text{RIT2FWNN}[e(t-1), e(t-2), u_s(t-1)]$$

Each three inputs of RIT2FWNN are fuzzified with two Gaussian type-2 MFs with a fixed center and uncertain standard deviation. The premise and consequent parameters of RIT2FWNN are tuned based on the proposed online SMC-based learning algorithm.

Figs. 3 and 4 show the control trajectory and tracking error, respectively. The trajectories of the states x_1 , y_1 and x_2 , y_2 are depicted in Fig. 5 that demonstrate synchronization is perfectly achieved. The three-dimensional phase portrait, synchronization performance, of the chaotic master and slave systems is shown in Fig. 6. It can be seen that the synchronization performance of the proposed method in the presence of external disturbances and is successfully realized and the control trajectory is also admissible and bounded. The mean square error (MSE) values for the synchronization errors (e_1 , e_2) are given in Table 3. It can be seen that the performances of the proposed method in this paper are significantly better than other approaches; (1) sliding mode technique [16], (2) non-linear model predictive

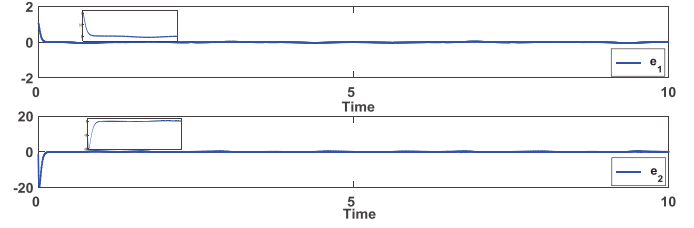


Fig. 3: Tracking error for Ex.2.

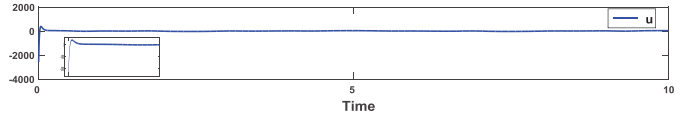


Fig. 4: Control signal for Ex.2.

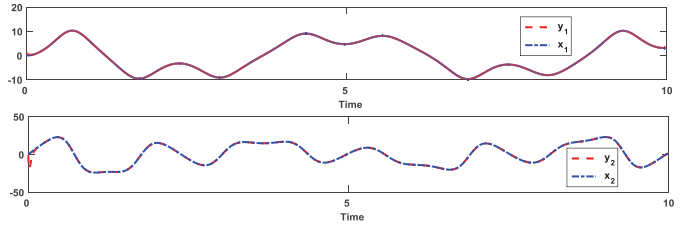


Fig. 5: Trajectories of the states x_1 , y_1 and x_2 , y_2 , Ex.2.

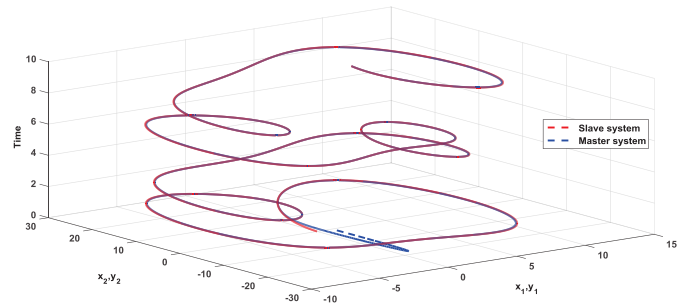


Fig. 6: 3D phase portrait, synchronization performance, of the master and slave systems for Ex.2.

controller (NMPC)[24]. Additionally, tracking error results in the presence of pulse disturbance are depicted in Fig. 7. As can be seen, the controller is robust against external disturbances.

Table 3: Comparison the values of MSE for different methods in Ex.2.

	SMC [16]	method	NMPC [24]	proposed method
e_1	2500		0.0031	0.0012
e_2	1900		1.2897	0.4333

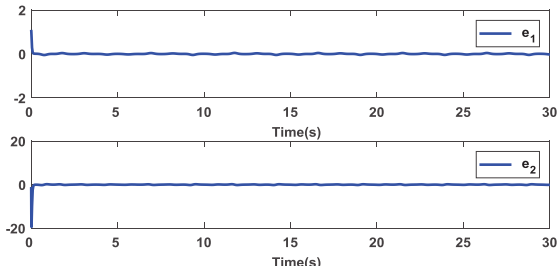


Fig. 7: Tracking error when the pulse disturbance applied for Ex.2.

Example 3: In this example, we will apply the proposed method to synchronize two nonidentical fractional-order time-delay chaotic systems. The master system is as follows [32]:

$$\begin{cases} D^\alpha y_1(t) = y_3 - 3y_1(t)y_2(t - 0.001) + \omega_1^m(t) \\ D^\alpha y_2(t) = 1 - 0.1y_2(t) - y_1^2(t - 0.001) + \omega_2^m(t) \\ D^\alpha y_3(t) = -y_1(t - 0.001) - y_3(t) + \omega_3^m(t) \end{cases} \quad (48)$$

The initial conditions are $y(0) = [y_1(0), y_2(0), y_3(0)]^T = [0.1, 4, 0.5]$ and the simulation sample time is 0.001. The slave system is Liu fractional-order time-delay chaotic system which its dynamics are as follows[32]:

$$\begin{cases} D^\alpha x_1(t) = 10(x_2(t) - x_1(t)) + \omega_1^s(t) + u_1(t) \\ D^\alpha x_2(t) = 40x_1(t - 0.001) - x_1(t)x_3(t) + \omega_2^s(t) + u_2(t) \\ D^\alpha x_3(t) = -2.5x_3(t - 0.001) + 4x_1^2(t) + \omega_3^s(t) + u_3(t) \end{cases} \quad (49)$$

where the fractional derivative order of both slave and master chaotic systems are $\alpha = 0.92$. The initial conditions are $[x_1(0), x_2(0), x_3(0)]^T = [0.1, 4, 0.5]$. The controller design procedure is the same as Example 2. $d_k^m(t)$ and $d_k^s(t), k = 1, 2, 3$ are the external bounded disturbances of slave and master system respectively which are initialized as follows:

$$\begin{aligned} \omega^m(t) &= (\omega_1^m(t), \omega_2^m(t), \omega_3^m(t))^T \\ &= (-0.1\cos(10t), 0.2\cos(20t), 0.1\sin(10t))^T \\ \omega^s(t) &= (\omega_1^s(t), \omega_2^s(t), \omega_3^s(t))^T \\ &= (0.1\sin(20t), -0.3\sin(10t), -0.5\cos(10t))^T \end{aligned} \quad (50)$$

The trajectories of the states x_k, y_k where $k = 1, 2, 3$ and the control signals are shown in Figs.8 and 9, respectively. As one can see from the figure, the tracking performance is good even in the presence of disturbances and unknown functions of dynamics of the system. The designed controller can synchronize effectively two non-identical fractional-order time-delay chaotic systems.

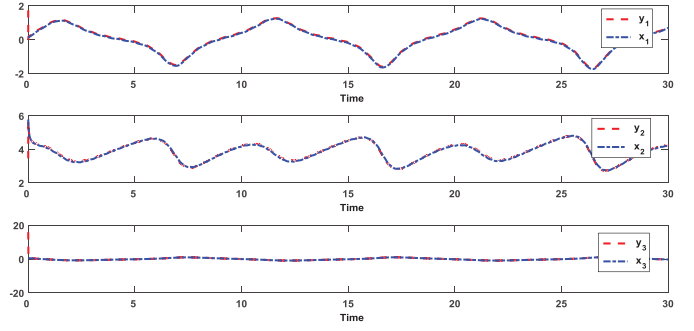


Fig. 8: Trajectories of the states x_k, y_k for $k = 1, 2, 3$ for Ex.3

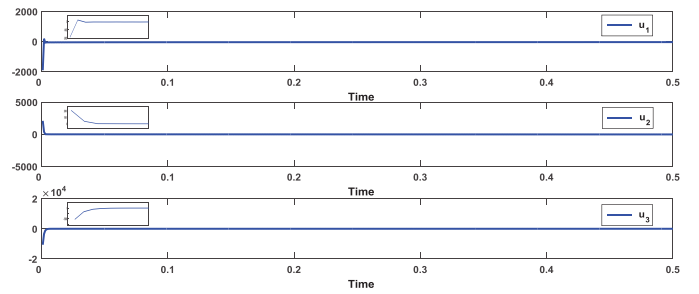


Fig. 9: Control signals for Ex.3

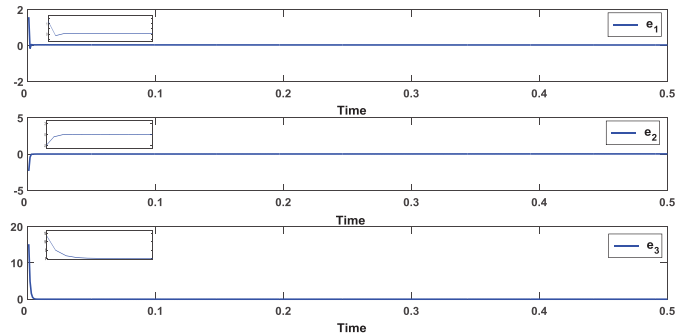


Fig. 10: Tracking error for Ex.3.

From Fig.10, it can be observed that the tracking errors $e_1(t), e_2(t)$ and $e_3(t)$ converge both to small intervals around zero. Table 4 indicates that the obtained MSE results are comparable with the results in [32]. Fig. 11 shows the tracking error results in the presence of pulse disturbances. It is shown that the developed controller preserves the robustness properties against external disturbances.

6 Conclusion and Future Work

In this paper, a novel robust model RIT2FWNN with fully SMC-based learning algorithm is proposed to design a model-based predictive controller to deal with

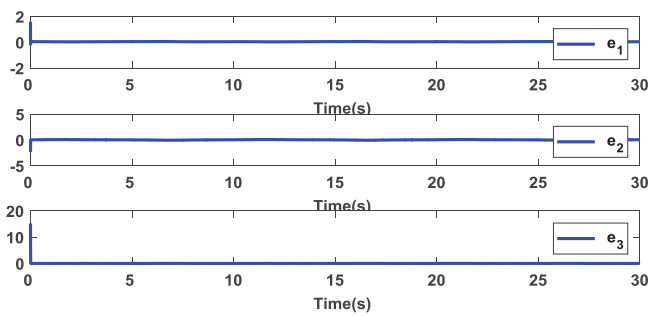


Fig. 11: Tracking error when the pulse disturbance applied for Ex.3.

Table 4: Comparison the values of MSE for different methods in Ex.3.

	<i>Hybrid projective synchronization</i>	<i>proposed method</i>
e_1	0.15	0.019
e_2	0.95	0.021
e_3	1.43	0.086

synchronization between two different fractional-order time-delay chaotic systems with subject to uncertainty and external additive disturbances. This synchronization is effectively obtained which demonstrate the superior performance of the proposed methodology. The model is tested against identification of a nonlinear and time-varying dynamic system. Since the output of these system depends on different step delayed outputs and inputs, the task of identification of such systems is difficult. The simulation results indicate that the proposed approach is quite useful in modeling unknown function of dynamic systems, having not only favorable tracking performance but also produce strong robustness and faster convergence speed. The Lyapunov stability theorem is applied to derive the parameter update rule of the model to guarantee all parameters uniformly ultimately bounded.

The proposed RIT2FWNN owns some advantages:

1. The proposed SMC-based learning strategy efficiently updates the parameters of RIT2FWNN to cope with on-line problems which face uncertainty and external disturbances. Additionally, the adaptive learning rate can get fast convergence for the systems.
2. The parameters adjustments can guarantee the convergence of RIT2FWNN that can be maintained in experimental results and theoretical analysis.
3. Utilizing wavelet function and considering recurrent part in the structure along side with the proposed learning method make the performance of the proposed model better with respect to other learning

methods under uncertain conditions in the presence of disturbances.

4. High Synchronization accuracy which implies that the proposed model is a suitable choice. Hence, it is applied to MPC applications.

However the proposed RIT2FWNN network has disadvantages such as fixed structure which makes it complex spatially for systems with higher numbers of inputs.

As future work, there are many challenging issues remaining to investigate, including the choice of higher order sliding mode control in the process of training, employing self-structuring methods to achieve optimized construction of RIT2FWNN network.

References

1. Arena, P., Caponetto, R., Fortuna, L., Porto, D.: Chaos in a fractional order duffing system. In: Proceedings EC-CTD, Budapest, vol. 1259, p. 1262 (1997)
2. Azar, A.T., Serranot, F.E., Vaidyanathan, S.: Sliding mode stabilization and synchronization of fractional order complex chaotic and hyperchaotic systems. In: Mathematical Techniques of Fractional Order Systems, pp. 283–317. Elsevier (2018)
3. Begian, M.B., Melek, W.W., Mendel, J.M.: Parametric design of stable type-2 tsf fuzzy systems. In: Fuzzy Information Processing Society, 2008. NAFIPS 2008. Annual Meeting of the North American, pp. 1–6. IEEE (2008)
4. Boukroune, A., Bouzeriba, A., Bouden, T.: Fuzzy generalized projective synchronization of incommensurate fractional-order chaotic systems. *Neurocomputing* **173**, 606–614 (2016)
5. Cheng, R., Bai, Y.: A novel approach to fuzzy wavelet neural network modeling and optimization. *International Journal of Electrical Power & Energy Systems* **64**, 671–678 (2015)
6. Duan, F., Dai, L., Chang, W., Chen, Z., Zhu, C., Li, W.: semg-based identification of hand motion commands using wavelet neural network combined with discrete wavelet transform. *IEEE Transactions on Industrial Electronics* **63**(3), 1923–1934 (2016)
7. Ginarsa, I., Soeprijanto, A., Purnomo, M.: Controlling chaos and voltage collapse using an anfis-based composite controller-static var compensator in power systems. *International Journal of Electrical Power & Energy Systems* **46**, 79–88 (2013)
8. Han, H.G., Chen, Z.Y., Liu, H.X., Qiao, J.F.: A self-organizing interval type-2 fuzzy-neural-network for modeling nonlinear systems. *Neurocomputing* **290**, 196–207 (2018)
9. Jain, L.C., Seera, M., Lim, C.P., Balasubramaniam, P.: A review of online learning in supervised neural networks. *Neural computing and applications* **25**(3-4), 491–509 (2014)
10. Juang, C.F., Hsieh, C.D.: A locally recurrent fuzzy neural network with support vector regression for dynamic-system modeling. *IEEE Transactions on Fuzzy Systems* **18**(2), 261–273 (2010)
11. Juang, C.F., Huang, R.B., Lin, Y.Y.: A recurrent self-evolving interval type-2 fuzzy neural network for dynamic

- system processing. *IEEE Transactions on Fuzzy Systems* **17**(5), 1092–1105 (2009)
12. Kayacan, E., Kayacan, E., Khanesar, M.A.: Identification of nonlinear dynamic systems using type-2 fuzzy neural networks—a novel learning algorithm and a comparative study. *IEEE Transactions on Industrial Electronics* **62**(3), 1716–1724 (2015)
 13. Khazaei, M., Sadat-Hosseini, H., Marjaninejad, A., Daneshvar, S.: A radial basis function neural network approximator with fast terminal sliding mode-based learning algorithm and its application in control systems. In: *Electrical Engineering (ICEE), 2017 Iranian Conference on*, pp. 812–816. IEEE (2017)
 14. Khettab, K., Bensafia, Y., Bourouba, B., Azar, A.T.: Enhanced fractional order indirect fuzzy adaptive synchronization of uncertain fractional chaotic systems based on the variable structure control: Robust h8 design approach. In: *Mathematical Techniques of Fractional Order Systems*, pp. 597–624. Elsevier (2018)
 15. Lee, C.H., Teng, C.C.: Identification and control of dynamic systems using recurrent fuzzy neural networks. *IEEE Transactions on fuzzy systems* **8**(4), 349–366 (2000)
 16. Lin, T.C., Lee, T.Y.: Chaos synchronization of uncertain fractional-order chaotic systems with time delay based on adaptive fuzzy sliding mode control. *IEEE Transactions on Fuzzy Systems* **19**(4), 623–635 (2011)
 17. Lin, Y.Y., Chang, J.Y., Lin, C.T.: Identification and prediction of dynamic systems using an interactively recurrent self-evolving fuzzy neural network. *IEEE transactions on neural networks and learning systems* **24**(2), 310–321 (2013)
 18. Lin, Y.Y., Chang, J.Y., Pal, N.R., Lin, C.T.: A mutually recurrent interval type-2 neural fuzzy system (mrit2nfs) with self-evolving structure and parameters. *IEEE Transactions on Fuzzy Systems* **21**(3), 492–509 (2013)
 19. Lin, Y.Y., Chang, J.Y., Pal, N.R., Lin, C.T.: A mutually recurrent interval type-2 neural fuzzy system (mrit2nfs) with self-evolving structure and parameters. *IEEE Transactions on Fuzzy Systems* **21**(3), 492–509 (2013)
 20. Linhares, L.L., Fontes, A.I., Martins, A.M., Araújo, F.M., Silveira, L.F.: Fuzzy wavelet neural network using a coreentropy criterion for nonlinear system identification. *Mathematical Problems in Engineering* **2015** (2015)
 21. Lu, C.H.: Wavelet fuzzy neural networks for identification and predictive control of dynamic systems. *IEEE Transactions on Industrial Electronics* **58**(7), 3046–3058 (2011)
 22. Masumppoor, S., Khanesar, M.A., et al.: Adaptive sliding-mode type-2 neuro-fuzzy control of an induction motor. *Expert Systems with Applications* **42**(19), 6635–6647 (2015)
 23. Mohadeszadeh, M., Delavari, H.: Synchronization of uncertain fractional-order hyper-chaotic systems via a novel adaptive interval type-2 fuzzy active sliding mode controller. *International Journal of Dynamics and Control* **5**(1), 135–144 (2017)
 24. Mohammadzadeh, A., Ghaemi, S., Kaynak, O., et al.: Robust predictive synchronization of uncertain fractional-order time-delayed chaotic systems. *Soft Computing* pp. 1–16 (2018)
 25. Oviedo, J.J.E., Vandewalle, J.P., Wertz, V.: *Fuzzy logic, identification and predictive control*. Springer Science & Business Media (2006)
 26. Pan, Y., Yang, G.H.: Event-triggered fault detection filter design for nonlinear networked systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* **48**(11), 1851–1862 (2017)
 27. Pan, Y., Yang, G.H.: Event-triggered fuzzy control for nonlinear networked control systems. *Fuzzy Sets and Systems* **329**, 91–107 (2017)
 28. Petras, I.: A note on the fractional-order cellular neural networks. In: *Neural Networks, 2006. IJCNN'06. International Joint Conference on*, pp. 1021–1024. IEEE (2006)
 29. Podlubny, I.: Fractional-order systems and pi/sup/spl lambda/d/sup/spl mu//-controllers. *IEEE Transactions on automatic control* **44**(1), 208–214 (1999)
 30. Sira-Ramirez, H., Colina-Morles, E.: A sliding mode strategy for adaptive learning in adalines. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* **42**(12), 1001–1012 (1995)
 31. Unbehauen, H.: *Control Systems, Robotics and Automation*. Eolss Publishers Company Limited (2009)
 32. Wang, S., Yu, Y., Wen, G.: Hybrid projective synchronization of time-delayed fractional order chaotic systems. *Nonlinear Analysis: Hybrid Systems* **11**, 129–138 (2014)
 33. Yu, X., Kaynak, O.: Sliding mode control with soft computing: A survey. *IEEE transactions on industrial electronics* **56**(9), 3275–3285 (2009)
 34. Zheng, D.D., Xie, W.F., Chai, T., Fu, Z.: Identification and trajectory tracking control of nonlinear singularly perturbed systems. *IEEE Transactions on Industrial Electronics* **64**(5), 3737–3747 (2017)
- [PROOF OF THEOREM 1]
The time derivative of (16) is calculated as follows:
- $$\dot{\underline{w}}_r = \frac{\dot{\underline{w}}_r \sum_{r=1}^N \underline{w}_r - \sum_{r=1}^N \dot{\underline{w}}_r \underline{w}_r}{\left(\sum_{r=1}^N \underline{w}_r\right)^2} = \frac{\dot{\underline{w}}_r - \tilde{\underline{w}}_r \sum_{r=1}^N \dot{\underline{w}}_r}{\sum_{r=1}^N \underline{w}_r} \quad (51)$$
- $$\dot{\bar{w}}_r = \frac{\dot{\bar{w}}_r \sum_{r=1}^N \bar{w}_r - \sum_{r=1}^N \dot{\bar{w}}_r \bar{w}_r}{\left(\sum_{r=1}^N \bar{w}_r\right)^2} = \frac{\dot{\bar{w}}_r - \tilde{\bar{w}}_r \sum_{r=1}^N \dot{\bar{w}}_r}{\sum_{r=1}^N \bar{w}_r}$$
- considering the following equations for $\dot{\underline{w}}_r$ and $\dot{\bar{w}}_r$:
- $$\dot{\underline{w}}_r = \underline{J}_r \underline{w}_r; \quad \dot{\bar{w}}_r = \bar{J}_r \bar{w}_r \quad (52)$$
- thus the equation (51) changes as follows:
- $$\dot{\underline{w}}_r = -\tilde{\underline{w}}_r \bar{J}_r + \tilde{\underline{w}}_r \sum_{r=1}^N \tilde{\underline{w}}_r \bar{J}_r; \quad \dot{\bar{w}}_r = -\tilde{\bar{w}}_r \underline{J}_r + \tilde{\bar{w}}_r \sum_{r=1}^N \tilde{\bar{w}}_r \underline{J}_r \quad (53)$$
- Furthermore, the time derivative of upper and lower of type-2 MFs are calculated as (54):
- $$\dot{\underline{\mu}}_{ik}(x_i) = -\left(\frac{(x_i + \xi_{ik} - c_{ik})(\dot{x}_i + \dot{\xi}_{ik} - \dot{c}_{ik})}{\underline{\sigma}_{ik}^2} - \frac{\dot{\underline{\sigma}}_{ik}(x_i + \xi_{ik} - c_{ik})^2}{\underline{\sigma}_{ik}^3}\right) \quad (54)$$
- $$\dot{\bar{\mu}}_{ik}(x_i) = -\left(\frac{(x_i + \xi_{ik} - c_{ik})(\dot{x}_i + \dot{\xi}_{ik} - \dot{c}_{ik})}{\bar{\sigma}_{ik}^2} - \frac{\dot{\bar{\sigma}}_{ik}(x_i + \xi_{ik} - c_{ik})^2}{\bar{\sigma}_{ik}^3}\right)$$
- and the time derivative of the recurrent parameter is as follows:
- $$\dot{\xi}_{ik} = \frac{\dot{\theta}_{ik}(\underline{\mu}_{ik}(t-1) - \xi_{ik}) + \dot{\bar{\theta}}_{ik}(\bar{\mu}_{ik}(t-1) - \xi_{ik})}{(\theta_{ik} + \bar{\theta}_{ik})} \quad (55)$$
- considering $\dot{\underline{\mu}}_{ik}(x_i) = -\underline{A}_{ik} \dot{\underline{A}}_{ik}$ where \underline{A}_{ik} and $\dot{\underline{A}}_{ik}$ are as follows:

$$\begin{aligned} \underline{A}_{ik} &= \frac{x_i}{\underline{\sigma}_{ik}} - \frac{c_{ik}}{\underline{\sigma}_{ik}} + \frac{\underline{\theta}_{ik}\underline{\mu}_{ik}(t-1) + \bar{\theta}_{ik}\bar{\mu}_{ik}(t-1)}{\underline{\sigma}_{ik}(\underline{\theta}_{ik} + \bar{\theta}_{ik})} \\ \dot{\underline{A}}_{ik} &= \frac{\dot{x}_i}{\underline{\sigma}_{ik}} - \frac{\dot{c}_{ik}}{\underline{\sigma}_{ik}} - \dot{\underline{\sigma}}_{ik} \left(\frac{x_i - c_{ik} + \xi_{ik}}{\underline{\sigma}_{ik}^2} \right) + \frac{\dot{\underline{\theta}}_{ik}(\underline{\mu}_{ik}(t-1) - \xi_{ik})}{\underline{\sigma}_{ik}(\underline{\theta}_{ik} + \bar{\theta}_{ik})} + \frac{\dot{\bar{\theta}}_{ik}(\bar{\mu}_{ik}(t-1) - \xi_{ik})}{\underline{\sigma}_{ik}(\underline{\theta}_{ik} + \bar{\theta}_{ik})} \end{aligned} \quad (56)$$

By using Maclaurin series expansion:

$$\begin{aligned} \frac{1}{\underline{\sigma}_{ik}} &= \frac{1}{\frac{\underline{\sigma}_{ik} + \bar{\sigma}_{ik}}{2} + \frac{\underline{\sigma}_{ik} - \bar{\sigma}_{ik}}{2}} = \frac{2}{\bar{\sigma}_{ik} + \underline{\sigma}_{ik}} \underbrace{\left(1 - \frac{\underline{\sigma}_{ik} - \bar{\sigma}_{ik}}{\bar{\sigma}_{ik} + \underline{\sigma}_{ik}} + \left(\frac{\bar{\sigma}_{ik} - \underline{\sigma}_{ik}}{\bar{\sigma}_{ik} + \underline{\sigma}_{ik}} \right)^2 + H.O.T \right)}_{\underline{D}_{ik}} \\ \frac{1}{\bar{\sigma}_{ik}} &= \frac{1}{\frac{\bar{\sigma}_{ik} + \underline{\sigma}_{ik}}{2} + \frac{\bar{\sigma}_{ik} - \underline{\sigma}_{ik}}{2}} = \frac{2}{\underline{\sigma}_{ik} + \bar{\sigma}_{ik}} \underbrace{\left(1 - \frac{\bar{\sigma}_{ik} - \underline{\sigma}_{ik}}{\underline{\sigma}_{ik} + \bar{\sigma}_{ik}} + \left(\frac{\underline{\sigma}_{ik} - \bar{\sigma}_{ik}}{\underline{\sigma}_{ik} + \bar{\sigma}_{ik}} \right)^2 + H.O.T \right)}_{\bar{D}_{ik}} \end{aligned} \quad (57)$$

\underline{D}_{ik} and \bar{D}_{ik} are limited as follows:

$$(1 + \underline{D}_{ik})^2 + 1 < B_D, \quad (1 + \bar{D}_{ik})^2 + 1 < B_D \quad (58)$$

thus the \bar{J}_r and \underline{J}_r in (53) are as follows:

$$\bar{J}_r = - \sum_{i=1}^I \bar{A}_{ik} \dot{\bar{A}}_{ik}; \quad \underline{J}_r = - \sum_{i=1}^I \underline{A}_{ik} \dot{\underline{A}}_{ik} \quad (59)$$

by substituting (19)-(23) in the above equation:

$$\bar{J}_r = \underline{J}_r = I \alpha \text{sgn}(e) \quad (60)$$

In order to check the stability condition, the following Lyapunov function is considered:

$$V = \frac{1}{2} e^2 + \frac{1}{2\gamma} (\alpha - \alpha^*)^2 \quad (61)$$

The stability condition is met when the time derivative of V is negative definite ($\dot{V} < 0$). So,

$$\dot{V} = \dot{e}e + \frac{1}{\gamma} \dot{\alpha}(\alpha - \alpha^*) = e(y_N - \dot{y}) + \frac{1}{\gamma} \dot{\alpha}(\alpha - \alpha^*) \quad (62)$$

by differentiating (17)

$$\dot{y}_N = \dot{q} \sum_{r=1}^N f_r \underline{\tilde{w}}_r + q \sum_{r=1}^N \left(\dot{f}_r \underline{\tilde{w}}_r + f_r \dot{\underline{\tilde{w}}}_r \right) - \dot{q} \sum_{r=1}^N f_r \tilde{w}_r + (1-q) \sum_{r=1}^N \left(\dot{f}_r \tilde{w}_r + f_r \dot{\tilde{w}}_r \right) \quad (63)$$

by substituting (53):

$$\dot{y}_N = \dot{q} \sum_{r=1}^N f_r \underline{\tilde{w}}_r + q \sum_{r=1}^N \left(\dot{f}_r \underline{\tilde{w}}_r + f_r (-\underline{\tilde{w}}_r \underline{J}_r + \underline{\tilde{w}}_r \sum_{r=1}^N \underline{\tilde{w}}_r \underline{J}_r) \right) - \dot{q} \sum_{r=1}^N f_r \tilde{w}_r + (1-q) \sum_{r=1}^N \left(\dot{f}_r \tilde{w}_r + f_r (-\tilde{w}_r \bar{J}_r + \tilde{w}_r \sum_{r=1}^N \tilde{w}_r \bar{J}_r) \right) \quad (64)$$

by considering (60):

$$\dot{y}_N = \dot{q} \sum_{r=1}^N f_r \underline{\tilde{w}}_r + q \sum_{r=1}^N \left(\dot{f}_r \underline{\tilde{w}}_r - I \alpha \text{sgn}(e) f_r (\underline{\tilde{w}}_r - \underline{\tilde{w}}_r \sum_{r=1}^N \underline{\tilde{w}}_r) \right) - \dot{q} \sum_{r=1}^N f_r \tilde{w}_r + (1-q) \sum_{r=1}^N \left(\dot{f}_r \tilde{w}_r - I \alpha \text{sgn}(e) f_r (\tilde{w}_r - \tilde{w}_r \sum_{r=1}^N \tilde{w}_r) \right) \quad (65)$$

since $\sum_{r=1}^N \tilde{w}_r = 1$ and $\sum_{r=1}^N \underline{\tilde{w}}_r = 1$ equation (65) can be summarized and then by substituting \dot{q} according to (27),

$$\dot{y}_N = - \frac{1}{F(\underline{W} - \tilde{W})^T} \alpha \text{sgn}(e) \sum_{r=1}^N f_r (\underline{\tilde{w}}_r - \tilde{w}_r) + \sum_{r=1}^N \dot{f}_r (q \underline{\tilde{w}}_r + (1-q) \tilde{w}_r) = -\alpha \text{sgn}(e) + \sum_{r=1}^N \dot{f}_r (q \underline{\tilde{w}}_r + (1-q) \tilde{w}_r) \quad (66)$$

and the time derivative of f_r in (10) is calculated as:

$$\dot{f}_r = \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \left(-\frac{1}{2} \dot{a}_{ri} a_{ri}^{-1} |a_{ri}|^{-\frac{1}{2}} (1 - z_{ri}^2) e^{-\frac{z_{ri}^2}{2}} \right) + \rho_{ri} \left(|a_{ri}|^{-\frac{1}{2}} \left[-2 \dot{z}_{ri} z_{ri} e^{-\frac{z_{ri}^2}{2}} - \dot{z}_{ri} z_{ri} e^{-\frac{z_{ri}^2}{2}} (1 - z_{ri}^2) \right] \right) \quad (67)$$

$$\dot{f}_r = \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \left(-\frac{1}{2} \dot{a}_{ri} a_{ri}^{-1}\right) \Psi_{ri}(z) - \sum_{i=1}^I \rho_{ri} \dot{z}_{ri} z_{ri} \left[2 \frac{\Psi_{ri}(z)}{1 - z_{ri}^2} + \Psi_{ri}(z)\right] \quad (68)$$

$$\begin{aligned} &= \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \Psi_{ri}(z) \left(-\frac{1}{2} \dot{a}_{ri} a_{ri}^{-1} - \dot{z}_{ri} z_{ri} \left(\frac{2}{1 - z_{ri}^2} + 1\right)\right) \\ &= \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \Psi_{ri}(z) \left(-\frac{1}{2} \dot{a}_{ri} a_{ri}^{-1} - \dot{z}_{ri} z_{ri} \left(\frac{3 - z_{ri}^2}{1 - z_{ri}^2}\right)\right) \\ &= \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \Psi_{ri}(z) \left(-\frac{1}{2} \dot{a}_{ri} a_{ri}^{-1} - \left(\frac{\dot{x}_i - \dot{b}_{ri}}{a_{ri}} - \dot{a}_{ri} \frac{x_i - b_{ri}}{a_{ri}^2}\right) z_{ri} \left(\frac{3 - z_{ri}^2}{1 - z_{ri}^2}\right)\right) \end{aligned}$$

$$\dot{f}_r = \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \Psi_{ri}(z) \left(\dot{a}_{ri} \left(-\frac{1}{2} a_{ri}^{-1} + \frac{x_i - b_{ri}}{a_{ri}^2} \frac{3 - z_{ri}^2}{1 - z_{ri}^2} z_{ri}\right)\right) - \sum_{i=1}^I \rho_{ri} \Psi_{ri}(z) \left(\frac{\dot{x}_i}{a_{ri}} \frac{3 - z_{ri}^2}{1 - z_{ri}^2} z_{ri} + \frac{\dot{b}_{ri}}{a_{ri}} z_{ri} \left(\frac{3 - z_{ri}^2}{1 - z_{ri}^2}\right)\right) \quad (69)$$

so \dot{y}_N in (66) changes as follows:

$$\dot{y}_N = -\alpha \operatorname{sgn}(e) + \sum_{r=1}^N \sum_{i=1}^I \dot{\rho}_{ri} \Psi_{ri}(z) + \sum_{i=1}^I \rho_{ri} \Psi_{ri}(z) \left(\dot{a}_{ri} \left(-\frac{1}{2} a_{ri}^{-1} + \frac{x_i - b_{ri}}{a_{ri}^2} \frac{3 - z_{ri}^2}{1 - z_{ri}^2} z_{ri}\right) - \frac{\dot{x}_i}{a_{ri}} \frac{3 - z_{ri}^2}{1 - z_{ri}^2} z_{ri} + \frac{\dot{b}_{ri}}{a_{ri}} z_{ri} \left(\frac{3 - z_{ri}^2}{1 - z_{ri}^2}\right)\right) (q\tilde{w}_r + (1 - q)\tilde{w}_r) \quad (70)$$

substituting \dot{y}_N (70) and then \dot{a}_{ri} , \dot{b}_{ri} and $\dot{\rho}_{ri}$ according to (24) to (26) into the time derivative of Lyapunov function (62):

$$\dot{V} = -|e|(3IN + 1)\alpha + e \left[\sum_{r=1}^N \left[\sum_{i=1}^I \rho_{ri} (-\Psi_{ri}(z)) \frac{\dot{x}_i}{a_{ri}} \frac{3 - z_{ri}^2}{1 - z_{ri}^2} z_{ri} (q\tilde{w}_r + (1 - q)\tilde{w}_r) \right] - \dot{y} \right] + \frac{1}{\gamma} \dot{\alpha} (\alpha - \alpha^*) \quad (71)$$

$$\begin{aligned} \dot{V} &< -|e|(3IN + 1)\alpha + |e| \left[\sum_{r=1}^N \left[\sum_{i=1}^I B_\rho (-B_\psi \frac{B_{\dot{x}}}{a_{ri}} \left| \frac{3 - z_{ri}^2}{1 - z_{ri}^2} ||z_{ri}| \right|) (q\tilde{w}_r + (1 - q)\tilde{w}_r) \right] - B_{\dot{y}} \right] + \frac{1}{\gamma} \dot{\alpha} (\alpha - \alpha^*) \\ &< -|e|(3IN + 1)\alpha + |e| (NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}}) + \frac{1}{\gamma} \dot{\alpha} (\alpha - \alpha^*) \end{aligned}$$

where, $\alpha^* \geq \frac{(NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}})}{(3NI + 1)}$ which is regarded as an unknown parameter and is determined during the tuning of learning rate. Thus,

$$\begin{aligned} \dot{V} &\leq -|e|(3IN + 1)\alpha + |e| (NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}}) + |e|(3IN + 1)\alpha^* - |e|(3IN + 1)\alpha^* + \frac{1}{\gamma} \dot{\alpha} (\alpha - \alpha^*) \\ &\leq |e| (NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}}) - |e|(3IN + 1)\alpha^* - (\alpha - \alpha^*) \left((3IN + 1)|e| - \frac{1}{\gamma} \dot{\alpha} \right) \end{aligned} \quad (72)$$

the adaptive learning rate (α) changes as follows:

$$\dot{\alpha} = (3IN + 1)\gamma|e| - \nu\gamma\alpha \quad (73)$$

in which ν has small real value. The time derivative of Lyapunov function (72) is changed as follow:

$$\begin{aligned} \dot{V} &\leq |e| [NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}}] - |e|(3IN + 1)\alpha^* - (\alpha - \alpha^*)\nu\alpha \\ &\leq |e| [NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}}] - |e|(3IN + 1)\alpha^* - \nu \left(\alpha - \frac{\alpha^*}{2} \right)^2 + \frac{\nu\alpha^{*2}}{4} \end{aligned} \quad (74)$$

since $\alpha^* \geq \frac{(NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}})}{(3NI + 1)}$

$$|e| (NIB_\rho B_\psi B_{\dot{x}} B_a - B_{\dot{y}}) - |e|(3IN + 1)\left(\frac{\alpha^*}{2}\right) \leq 0 \quad (75)$$

consequently,

$$\dot{V} \leq -\left(\frac{\alpha^*}{2}\right)(3IN + 1)|e| + \frac{\nu\alpha^{*2}}{4} \quad (76)$$

In this case in order to have a negative definite time derivative of the Lyapunov function, it is required that the following condition holds.

$$|e| \leq \frac{\alpha^* \nu}{2(3IN + 1)} \quad (77)$$

Hence, it is concluded that error converges to a small neighborhood of zero, in which and stays there. Furthermore, the radius of this neighborhood can be made as small as desired using appropriate values for α^* and ν . This concludes the proof.