

Wavelet Neural Network Observer Based Adaptive Tracking Control for a Class of Uncertain Nonlinear Delayed Systems Using Reinforcement Learning

Manish Sharma

Dept. of Electronics and Instrumentation Engineering,
Medicaps Institute of Technology and Management, Indore, India
Email: er.mann24@gmail.com

Ajay Verma

Dept. of Electronics and Instrumentation Engineering,
Institute of Engineering and Technology, Devi Ahilya University, Indore, India

Abstract— This paper is concerned with the observer designing problem for a class of uncertain delayed nonlinear systems using reinforcement learning. Reinforcement learning is used via two Wavelet Neural networks (WNN), critic WNN and action WNN, which are combined to form an adaptive WNN controller. The “strategic” utility function is approximated by the critic WNN and is minimized by the action WNN. Adaptation laws are developed for the online tuning of wavelets parameters. By Lyapunov approach, the uniformly ultimate boundedness of the closed-loop tracking error is verified. Finally, a simulation example is shown to verify the effectiveness and performance of the proposed method.

Index Terms—Wavelet neural networks, adaptive control, optimal control, reinforcement learning, Lyapunov-Krasovskii functional.

I. INTRODUCTION

In many practical systems, the system model always contains some uncertain elements; these uncertainties may be due to additive unknown internal or external noise, environmental influence, nonlinearities such as hysteresis or friction, poor plant knowledge, reduced-order models, and uncertain or slowly varying parameters. Hence, the state observer for the uncertain system will be useful and apply to reconstruct the states of a dynamic system. The means to design adaptive observers through estimation of states and parameters in linear and nonlinear systems has been actively studied in recent years [1-3]. Adaptive observers of nonlinear systems have attracted much attention due to their wide uses in theory and practice.

The phenomenon of time-delay exists in various engineering systems such as chemical process, long transmission lines in pneumatic systems, etc. Time-delay usually leads to unsatisfactory performances and is frequently a source of instability. Recently, much effort

has been devoted to design the observer or observer-based control of time delay systems [4-6].

Reinforcement learning (RL) is a class of algorithms for solving multi-step, sequential decision problems by finding a policy for choosing sequences of actions that optimize the sum of some performance criterion over time [7-9]. In RL problems, an agent interacts with an unknown environment. At each time step, the agent observes the state, takes an action, and receives a reward. The goal of the agent is to learn a *policy* (i.e., a mapping from states to actions) that maximizes the long-term return. Actor-Critic algorithm is an implementation of RL which has separate structures for perception (critic) and action (actor) [10-12]. Given a specific state, the actor decides what action to take and the critic evaluates the outcome of the action in terms of future reward (goal).

Recently, wavelet neural networks (WNNs), which absorbs the advantages such as the multi-resolution of wavelets and the learning of NN, were proposed to guarantee the good convergence and were used to identify and control nonlinear systems [13-15]. Wavelet networks are feed-forward neural networks using wavelets as activation function. The WNN is suitable for the approximation of unknown nonlinear functions with local nonlinearities and fast variations because of its intrinsic properties of finite support and self-similarity.

Incorporating the advantages of WNN, adaptive actor-critic WNN-based control has emerged as a promising approach for the nonlinear systems. A reinforcement signal is used to evaluate the performance of a computational agent interacting with its environment. The aim is to discover a policy for selecting actions that minimize a specified performance index. Reinforcement learning methods are considered potentially useful to feedback control systems since these methods do not require comprehensive information about the system or its behaviour. During supervised learning with a neural network as the computational agent, the weights are adjusted in proportion to the output error of the neurons.

Something different occurs when using reinforcement learning – the weights are adjusted according to the reinforcement signal. This is especially useful since no desired output signals are available online to train a neural controller. The general idea of reinforcement learning is extended to neural networks using an adaptive critic model. In the actor-critic WNN based control; a long-term as well as short-term system-performance measure can be optimized. While the role of the actor is to select actions, the role of the critic is to evaluate the performance of the actor. This evaluation is used to provide the actor with a signal that allows it to improve its performance, typically by updating its parameters along an estimate of the gradient of some measure of performance, with respect to the actor's parameters. The critic WNN approximates a certain “strategic” utility function that is similar to a standard Bellman equation, which is taken as the long-term performance measure of the system. The weights of action WNN are tuned online by both the critic WNN signal and the filtered tracking error. It minimizes the strategic utility function and uncertain system dynamic estimation errors so that the optimal control signal can be generated. This optimal action NN control signal combined with an additional outer-loop conventional control signal is applied as the overall control input to the nonlinear system. The outer loop conventional signal allows the action and critic NNs to learn online while making the system stable. This conventional signal that uses the tracking error is viewed as the “supervisory” signal [7].

These motivate us to consider the designing of WNN observer based adaptive tracking controller for a class of uncertain delayed nonlinear systems using reinforcement learning. WNN are used for approximating the system uncertainty as well as to optimize the performance of the control strategy. The actor critic architecture based WNN is shown in figure 3.

The paper is organized as follows: section II deals with the system preliminaries, system description is given in section III. Section IV describes the wavelet observer design. WNN based controller designing aspects are discussed in section V. Section VI deals with the tuning algorithm for actor-critic wavelets. The stability analysis of the proposed control scheme and the observer is given in section VII. Effectiveness of the proposed strategy is illustrated through an example in section VIII while section IX concludes the paper.

II. SYSTEM PRELIMINARIES

A. Wavelet Neural Network

Wavelet network is a type of building block for function approximation. It is a single layer network consisting of translated and dilated versions of orthonormal father and mother wavelet function. Basis functions are used in wavelet network span $L^2(\mathfrak{R})$ subspace. Due to its universal approximation property any function $f(x) \in L^2(\mathfrak{R})$ can be approximated by linear combination of basis functions. The building block is obtained by translating and dilating the mother wavelet

function. Corresponding to certain countable family of a_m and b_n , wavelet function can be expressed as

$$\left\{ a_m^{-d/2} \psi \left(\frac{x-b_n}{a_m} \right) : m \in \mathbb{Z}, n \in \mathbb{Z}^d \right\} \quad (1)$$

Considering

$$a_m = a_0^m, b_n = n a_0^{-m} b_0, m \in \mathbb{Z}, n \in \mathbb{Z}^d \quad (2)$$

The wavelet in (1) can be expressed as

$$\psi_{mn} = \left\{ a_0^{-md/2} \psi \left(a_0^{-m} x - n b_0 \right) : m \in \mathbb{Z}, n \in \mathbb{Z}^d \right\}$$

Where the scalar parameters a_0 and b_0 define the step size of dilation and translation discretizations (typically $a_0 = 2$ and $b_0 = 1$) and $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ is the input vector.

Output of an n dimensional WNN with m wavelet nodes is

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}^d} \alpha_{mn} \psi_{mn} \quad (3)$$

B. System Description

Consider a nonlinear system of the form

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_n &= f[x(t), x(t-\tau)] + gu \\ y &= x_1 \end{aligned} \quad (4)$$

where $x = [x_1, x_2, \dots, x_n]^T$, u, y are state variable, control input and output respectively. τ is the known delay in the state. $f = [f_1, f_2, \dots, f_n]^T : \mathfrak{R}^{n+1} \rightarrow \mathfrak{R}^n$ are smooth unknown, nonlinear functions of state variables.

Rewriting the system (4) as

$$\dot{x} = Ax + B(f[x(t), x(t-\tau)] + gu) \quad (5)$$

$$y = Cx$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, C = [1 \ 0 \ 0 \ \dots \ 0]$$

The objective is to formulate a state feedback control law to achieve the desired tracking performance. The asymptotic tracking problem entails making a nonlinear system asymptotically track a desired reference trajectory while guaranteeing boundedness of all the internal variables. In its general formulation, even when the state of the plant is known, the tracking problem is considerably more difficult than that of stabilization. The main reason for its complexity lies in the fact that, even when the uncontrolled plant is an autonomous system, in order for the output to track the reference trajectory the controller should create and stabilize a time-varying equilibrium point in the closed-loop system which, in general, is difficult to calculate. The control law is formulated using the transformed system (5). Let

$\bar{y}_d = [y_d, \dot{y}_d, \dots, y_d^{(n-1)}]^T$ be the vector of desired tracking trajectory. Following assumptions are taken for the systems under consideration.

Assumption A1

- Desired trajectory $y_d(t)$ is assumed to be smooth, continuous C^n and available for measurement.
- For the system under consideration only output y is measurable.

III. WAVELET OBSERVER DESIGN

Wavelet based observer that estimates the states of the system(5) is given by

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + B[\hat{f}(\hat{x}(t), x(t-\tau)) + gu] - v_r + m(y - C\hat{x}) \quad (6) \\ \hat{y} &= C\hat{x} \end{aligned}$$

where \hat{x} is the estimation of state vector x , $m = [m_1, m_2, \dots, m_n]^T$ is the observer gain matrix, selected such that the matrix $A - mC$ is stable and v_r is the robust control term. In this work WNN is used for system identification.

Assumption A2

- $\|f(x(t), x(t-\tau)) - f(\hat{x}(t), \hat{x}(t-\tau))\| \leq \gamma_1 \|\tilde{x}\| + \gamma_2 \|\tilde{x}(t-\tau)\|$

- For a symmetric positive definite matrix Q there exist a symmetric positive definite matrix P such that

$$(A - mC)^T P + P(A - mC) = -Q \text{ and } (PB)^T = C$$

where $\tilde{x} = x - \hat{x}$ is the state variable estimation error while γ_1 and γ_2 are positive constants. Defining the error system as

$$\begin{aligned} \dot{\tilde{x}} &= (A - mC)\tilde{x} + B(f(x(t), x(t-\tau)) - f(\hat{x}(t), \hat{x}(t-\tau))) - v_r \quad (7) \\ \tilde{y} &= C\tilde{x} \end{aligned}$$

The robust control term is defined as

$$v_r = \frac{\tilde{y}(\rho^2 + 1)}{2\rho^2} \quad (8)$$

where ρ is the prescribed attenuation.

IV. BASIC CONTROLLER DESIGN USING FILTERED TRACKING ERROR

Defining the state tracking error vector $\hat{e}(t)$ as

$$\hat{e}(t) = \hat{x}(t) - \bar{y}_d(t)$$

The filter tracking error is defined as $\hat{r} = K\hat{e}$ where $K = [k_1, k_2, \dots, k_{n-1}, 1]$ is an appropriately chosen coefficient vector such that $\hat{e} \rightarrow 0$ exponentially as $\mathfrak{R} \rightarrow 0$.

Applying the feedback linearization method, the control law is defined as

$$u = \frac{1}{g} (y_d^{(n)} - k(\hat{f}(\hat{x}(t), \hat{x}(t-\tau)) + mC\tilde{x}) - K_e \hat{e} - k_r \hat{r}) \quad (9)$$

where $K_e = [0, k_1, k_2, \dots, k_{n-1}]$.

V. ADAPTIVE WNN CONTROLLER DESIGN

A novel strategic utility function is defined as the long-term performance measure for the system. It is approximated by the WNN critic signal. The action WNN signal is constructed to minimize this strategic utility function by using a quadratic optimization function. The critic WNN and action WNN weight tuning laws are derived. Stability analysis using the Lyapunov direct method is carried out for the closed-loop system (5) with novel weight tuning updates.

A. Strategic Utility Function

The utility function $p(k) = [p_i(k)]_{i=1}^m \in \mathfrak{R}^m$ is defined on the basis of the filtered tracking error \hat{r} and is given by:

$$\begin{aligned} p_i(k) &= 0 \quad \text{if } \hat{r}_i^2 \leq \eta \\ &= 1 \quad \text{if } \hat{r}_i^2 > \eta \end{aligned} \quad (10)$$

where $p_i(k) \in \mathfrak{R}$, $i = 1, 2, \dots, m$ and $\eta \in \mathfrak{R}$ is the predefined threshold. $p(k)$ can be considered as the current performance index defining good tracking performance for $p(k) = 0$ and poor tracking performance for $p(k) = 1$. The strategic utility function $Q'(k) \in \mathfrak{R}^m$ can be defined using the binary utility function as

$$Q'(k) = \alpha^N p(k+1) + \alpha^{N-1} p(k+2) + \dots + \alpha^{k+1} p(N) + \dots \quad (11)$$

where $\alpha \in \mathfrak{R}$ and $0 < \alpha < 1$ and N is the horizon. Above equation may be rewritten as

$$Q(k) = \min_{u(k)} \{ \alpha Q(k-1) - \alpha^{N+1} p(k) \} \quad (12)$$

B. Critic WNN

The long term system performance can be approximated by the critic WNN by defining the prediction error as

$$e_c(k) = \hat{Q}(k) - \alpha(\hat{Q}(k-1) - \alpha^N p(k)) \quad (13)$$

where $\hat{Q}(k) = \hat{w}_1^T(k) \phi_1(v_1^T x(k)) = \hat{w}_1^T(k) \phi_1(k)$, $e_c(k) \in \mathfrak{R}^m$, $\hat{Q}(k) \in \mathfrak{R}^m$ is the critic signal, $w_1(k) \in \mathfrak{R}^{n_1 \times m}$ and $v_1 \in \mathfrak{R}^{m \times n_1}$ represent the weight estimates, $\phi_1(k) \in \mathfrak{R}^{n_1}$ is the wavelet activation function and n_1 is the number of nodes in the wavelet layer. The objective function to be minimized by the critic NN is defined as:

$$E_c(k) = \frac{1}{2} e_c^T(k) e_c(k) \quad (14)$$

The weight update rule for the critic NN is derived from gradient-based adaptation, which is given by

$$\hat{w}_1(k+1) = \hat{w}_1(k) + \Delta \hat{w}_1(k) \quad (15)$$

where $\Delta \hat{w}_1(k) = \alpha_1 \left[-\frac{\partial E_c(k)}{\partial \hat{w}_1(k)} \right]$

Or

$$\hat{w}_1(k+1) = \hat{w}_1(k) - \alpha_1 \phi_1(k) \times (\hat{w}_1^T(k) \phi_1(k) + \alpha^{N+1} p(k) - \alpha \hat{w}_1^T(k-1) \phi_1(k-1))^T \quad (16)$$

where $\alpha_1 \in \mathfrak{R}$ is the WNN adaptation gain. The critic WNN weights are tuned by the reinforcement learning signal and the discounted past output values of critic WNN.

C. Action WNN

The action NN is implemented for the approximation of the unknown nonlinear function $f(x(k))$ and to provide an optimal control signal to the overall input $u(k)$ as

$$\hat{f}(k) = \hat{w}_2^T(k) \phi_2(v_2^T x(k)) = \hat{w}_2^T(k) \phi_2(k) \quad (17)$$

where $\hat{w}_2(k) \in \mathfrak{R}^{n_2 \times m}$ and $v_2 \in \mathfrak{R}^{m \times n_2}$ represent the matrix of weight estimate, $\phi_2(k) \in \mathfrak{R}^{n_2}$ is the activation function, n_2 is the number of nodes in the hidden layer. Suppose that the unknown target output-layer weight for the action WNN is w_2 then we have

$$f(k) = w_2^T(k) \phi_2(v_2^T x(k)) \varepsilon_2(x(k)) = w_2^T(k) \phi_2(k) \varepsilon_2(x(k)) \quad (18)$$

where $\varepsilon_2(x(k)) \in \mathfrak{R}^m$ is the WNN approximation error. From (17) and (18), we get

$$\tilde{f}(k) = \hat{f}(k) - f(k) = (\hat{w}_2(k) - w_2)^T \phi_2(k) - \varepsilon_2(x(k)) \quad (19)$$

where $\tilde{f}(k) \in \mathfrak{R}^m$ is the functional estimation error. The action WNN weights are tuned by using the functional estimation error $\tilde{f}(k)$ and the error between the desired strategic utility function $Q_d(k) \in \mathfrak{R}^m$ and the critic signal $\hat{Q}(k)$. Define

$$e_a(k) = \tilde{f}(k) + (\hat{Q}(k) - Q_d(k)) \quad (20)$$

The objective is to make the utility function $Q_d(k)$ zero at every step. Thus (20) becomes

$$e_a(k) = \tilde{f}(k) + \hat{Q}(k) \quad (21)$$

The objective function to be minimized by the action NN is given by

$$E_a(k) = \frac{1}{2} e_a^T(k) e_a(k) \quad (22)$$

The weight update rule for the action NN is also a gradient based adaptation, which is defined as

$$\hat{w}_2(k+1) = \hat{w}_2(k) + \Delta \hat{w}_2(k) \quad (23)$$

$$\text{where } \Delta \hat{w}_2(k) = \alpha_2 \left[-\frac{\partial E_a(k)}{\partial \hat{w}_2(k)} \right]$$

or

$$\hat{w}_2(k+1) = \hat{w}_2(k) - \alpha_2 \phi_2(k) (\hat{Q}(k) + \tilde{f}(k))^T \quad (24)$$

where $\alpha_2 \in \mathfrak{R}$ is the WNN adaptation gain.

The WNN weight updating rule in (24) cannot be implemented in practice since the nonlinear function

$f(x(k))$ is unknown. However, using (9), the functional estimation error is given by

$$\tilde{f}(k) = \dot{r} - r + \delta(k) \quad (25)$$

Substituting (25) in to (24), we get

$$\hat{w}_2(k+1) = \hat{w}_2(k) - \alpha_2 \phi_2(k) (\hat{Q}(k) + \dot{r} - r + \delta(k))^T$$

To implement the weight update rule, the unknown but bounded disturbance $\delta(k)$ is taken to be zero. Then, (25) is rewritten as

$$\hat{w}_2(k+1) = \hat{w}_2(k) - \alpha_2 \phi_2(k) (\hat{Q}(k) + \dot{r} - r)^T \quad (26)$$

Coincidentally, after replacing the functional approximation error, the weight update for the action NN is tuned by the critic WNN output, current filtered tracking error, and a conventional outer-loop signal. The block diagram of the overall closed loop system using actor critic WNN based observer is shown in figure 4.

VI. STABILITY ANALYSIS

Consider a Lyapunov-Krasovskii functional of the form

$$V = \frac{1}{2} \tilde{x}^T P \tilde{x} + \frac{1}{2} \hat{r}^2 + \int_{t-\tau}^t w(\sigma) d\sigma \quad (27)$$

Differentiating it along the trajectories of the system

$$\begin{aligned} \dot{V} = & -\frac{1}{2} \tilde{x}^T Q \tilde{x} + \tilde{x}^T P B (\delta(x, \bar{y}_d) - \hat{\delta}(\hat{x}, \bar{y}_d) - v_r) + \\ & \tilde{x}^T P (\phi(x(t), x(t-\tau)) - \phi(\hat{x}(t), \hat{x}(t-\tau))) \\ & + \hat{r} (K_e \hat{e} + K (\phi(\hat{x}(t), \hat{x}(t-\tau)) + m C \tilde{x}) + \hat{\delta}(\hat{x}, \bar{y}_d) \\ & + g v + v_r - y_d^n) + w(x(t)) - w(x(t-\tau)) \end{aligned}$$

$$\begin{aligned} \text{Substituting control law } v \text{ in the above equation} \\ = & -\frac{1}{2} Q_{\min} \|\tilde{x}^2\| + \tilde{x}^T P B (\delta(x, \bar{y}_d) - \delta(\hat{x}, \bar{y}_d) + \tilde{\delta}(\hat{x}, \bar{y}_d) - v_r) + \\ & \tilde{x}^T P (\phi(x(t), x(t-\tau)) - \phi(\hat{x}(t), \hat{x}(t-\tau))) + \hat{r} (v_r - k_r \hat{r}) + \\ & w(x(t)) - w(x(t-\tau)) \\ \leq & -\frac{1}{2} Q_{\min} \|\tilde{x}^2\| + M_1 \|\tilde{x}\|^2 + M_2 \|\tilde{x}\|^2 + M_3 \|\tilde{x}\| \|\tilde{x}(t-\tau)\| \\ & + \tilde{x}^T P B (\tilde{\delta}(\hat{x}, \bar{y}_d) - v_r) + M_4 + \hat{r} (-k_r \hat{r}) + w(x(t)) - w(x(t-\tau)) \end{aligned}$$

$$M_1 = \|PB\| \gamma_3, M_2 = P_{\max} \gamma_1, M_3 = P_{\max} \gamma_2, M_4 = \max |\hat{r} v_r| + \gamma_4$$

where $\gamma_1, \gamma_2, \gamma_3, \gamma_4 \geq 0$

Substituting the robust control term in above equation,

$$\begin{aligned} \leq & -\frac{1}{2} Q_{\min} \|\tilde{x}^2\| + M_1 \|\tilde{x}\|^2 + M_2 \|\tilde{x}\|^2 + M_3 \|\tilde{x}\| \|\tilde{x}(t-\tau)\| \\ & + \frac{\rho^2}{2} \varepsilon_\delta^2 - \frac{\tilde{y}^2}{2} + M_4 - k_r \hat{r}^2 + w(x(t)) - w(x(t-\tau)) \end{aligned}$$

Applying Young's inequality in the fourth term of the above equation,

$$M_3 \|\tilde{x}\| \|\tilde{x}(t-\tau)\| \leq \frac{M_3 \mu^2}{2} \|\tilde{x}\|^2 + \frac{M_3}{2\mu^2} \|\tilde{x}(t-\tau)\|^2$$

$$\text{Selecting } w(x(t)) = \frac{M_3}{2\mu^2} \|\tilde{x}\|^2 \text{ results in}$$

$$\begin{aligned} &\leq -\frac{1}{2}Q_{\min}\|\tilde{x}^2\| + (M_1 + M_2 + \frac{M_3\mu^2}{2} + \frac{M_3}{2\mu^2})\|\tilde{x}\|^2 \\ &+ \frac{\rho^2}{2}\varepsilon^2_{\delta} - \frac{\tilde{y}^2}{2} + M_4 - k_r\hat{r}^2 \end{aligned}$$

The system is stable as long as

$$\begin{aligned} k_r\hat{r}^2 + \frac{\tilde{y}^2}{2} + \frac{1}{2}Q_{\min}\|\tilde{x}^2\| &\geq (M_1 + M_2 + \frac{M_3\mu^2}{2} + \frac{M_3}{2\mu^2})\|\tilde{x}\|^2 \\ &+ \frac{\rho^2}{2}\varepsilon^2_{\delta} + M_4 \end{aligned} \quad (28)$$

VII. SIMULATION RESULTS

Simulation is performed to verify the effectiveness of proposed reinforcement learning WNN observer based control strategy. Considering a system of the form

$$\begin{aligned} \dot{x}_1 &= x_2(t) \\ \dot{x}_2 &= x_1(t)x_1(t-\tau)x_2(t-\tau)\sin x_2 + u \\ y &= x_1 \end{aligned} \quad (29)$$

System belongs to the class of uncertain nonlinear systems defined by (4) with $n=2$. The time delay τ is taken as 1 sec. It is assumed that only output is available for measurement. The proposed observer controller strategy is applied to this system with an objective to solve the tracking problem of system. The desired trajectory is taken as $y_d = 0.5\sin t + 0.1\cos 0.5t + 0.3$. Initial conditions are taken as $x(0) = [0.75, 0]^T$. Attenuation levels for robust controller are taken as 0.01. Controller gain vector is taken as $k = [25, 3]$. Wavelet networks with discrete Shannon's wavelet as the mother wavelet is used for approximating the unknown system dynamics. Wavelet parameters for these wavelet networks are tuned online using the proposed adaptation laws. Initial conditions for all the wavelet parameters are set to zero. Simulation results are shown in the figure. As observed from the figures, system response tracks the desired trajectory rapidly.

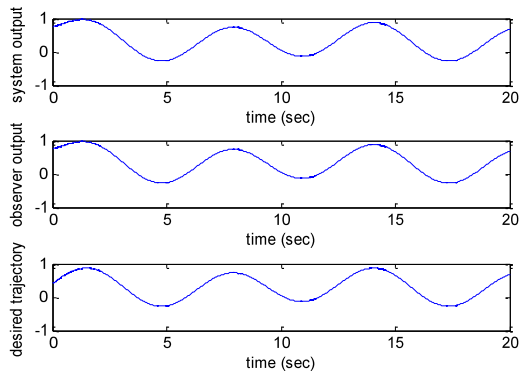


Figure 1. System output. Observer output and desired trajectory

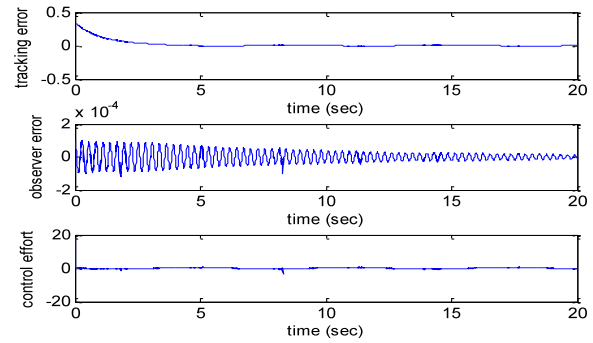


Figure 2. Tracking error, Observer error and control effort

VIII. CONCLUSION

A reinforcement learning WNN observer based adaptive tracking control strategy is proposed for a class of systems with unknown system dynamics. Adaptive wavelet networks are used for approximating the unknown system dynamics. Adaptation laws are developed for online tuning of the wavelet parameters. The stability of the overall system is guaranteed by using the Lyapunov functional. The theoretical analysis is validated by the simulation results. As observed from the figure 2 that tracking error rapidly converges to the small value of the order of 10^{-6} . A convergence pattern in the observer error is also reflected from the figure.

REFERENCES

- [1] H. Lens and J. Adamy, "Observer based controller design for linear systems with input constraints", Proceedings of 17th World Congress, The international federation of automatic control, Seoul Korea, pp.9916-9921, July 2008.
- [2] F. Abdollahi, H.A Talebi and R.V. Patel, "A stable neural network-based observer with application to flexible-joint manipulators", IEEE Transactions on Neural Networks, Vol. 17, Issue 1, , pp.118-129 Jan. 2006.
- [3] F. Sun; Z. Sun; P. Y. Woo, "Neural network-based adaptive controller design of robotic manipulators with an observer", IEEE Transactions on Neural Networks, Vol. 12, Issue 1, pp.54-67 Jan. 2001.
- [4] M. S. Mahmoud, Y. Shi and H. Nounou, "Resilient observer based control of uncertain time delay systems", International journal of Innovative Computing, Information and Control, Vol. 3 no.2, pp.407-418.Apr.2007.
- [5] P. Cui and C. Zhang, "Observer design and output feedback stabilization of linear singular time delay systems with unknown inputs". Journal of control theory application, Vol. 6, no.2, pp-177-183.Apr.2008.
- [6] J.P. Richard, "Time delay systems: an overview of some recent advances and open problems", Automatica, Vol. 39, pp. 1667-1694, 2003.
- [7] P. He and S. Jagannathan, "Reinforcement learning neural-network-based controller for nonlinear discrete-time systems with input constraints," IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, Vol. 37, no. 2, pp.425-436, April 2007.
- [8] W.S. Lin, L.H. Chang and P.C. Yang, "Adaptive critic anti-slip control of wheeled autonomous robot", IET Control Theory Applications, Vol. 1, No. 1, January 2007.

- [9] X. Lin and S. N. Balakrishnan, "Convergence analysis of adaptive critic based optimal control," Proceedings of the American Control Conferenc, pp. 1929–1933, 2000.
- [10] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs", IEEE Transactions on Neural Networks, vol. 8, no. 5, pp. 997–1007, Sep. 1997.
- [11] H. V. Hasselt and M. Wiering, "Reinforcement learning in continuous action spaces", IEEE Symposium on Approximate Dynamic Programming and Reinforcement Learning, pages 272–279, 2007.
- [12] Hussein Jaddu and Kunihiko Hiraishi, "A Wavelet Approach for Solving linear quadratic optimal control problems", SICE-ICASE International Joint Conference ,Oct. 18-21, Korea, 2006.
- [13] Q. Zhang and A. Benveniste, "Wavelet networks," IEEE Transactions on Neural Networks, Vol. 3, no. 6, pp.889-898, November 1992.
- [14] J. Zhang, G. G. Walter, Y. Miao, and W. Lee, "Wavelet neural networks for function learning," IEEE Transactions on Signal Processing, Vol. 43, no. 6, pp.1485-1497, June 1995.
- [15] B. Delyon, A. Juditsky, and A. Benveniste, "Accuracy analysis for wavelet approximations," IEEE Transactions on Neural Networks, Vol. 6, no. 2, pp.332-348 March 1995.

SHARMA Manish (1982 —) , male, Indore, India, Assistant Professor, Ph.D. candidate, his research directions include adaptive control and optimal control;
VERMA Ajay (1966 —) , male, Indore, India , Professor, supervisor for Ph.D. candidate, his research directions include signal processing, adaptive control and optimal control.

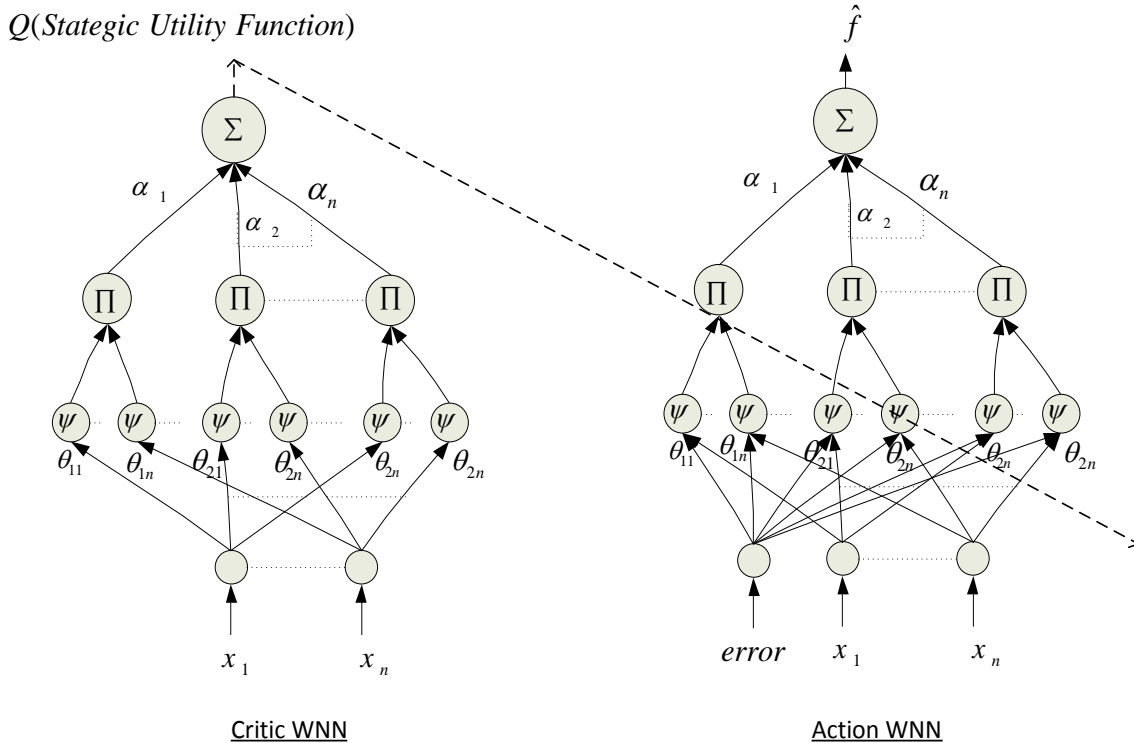


Figure 3. Actor Critic Architecture

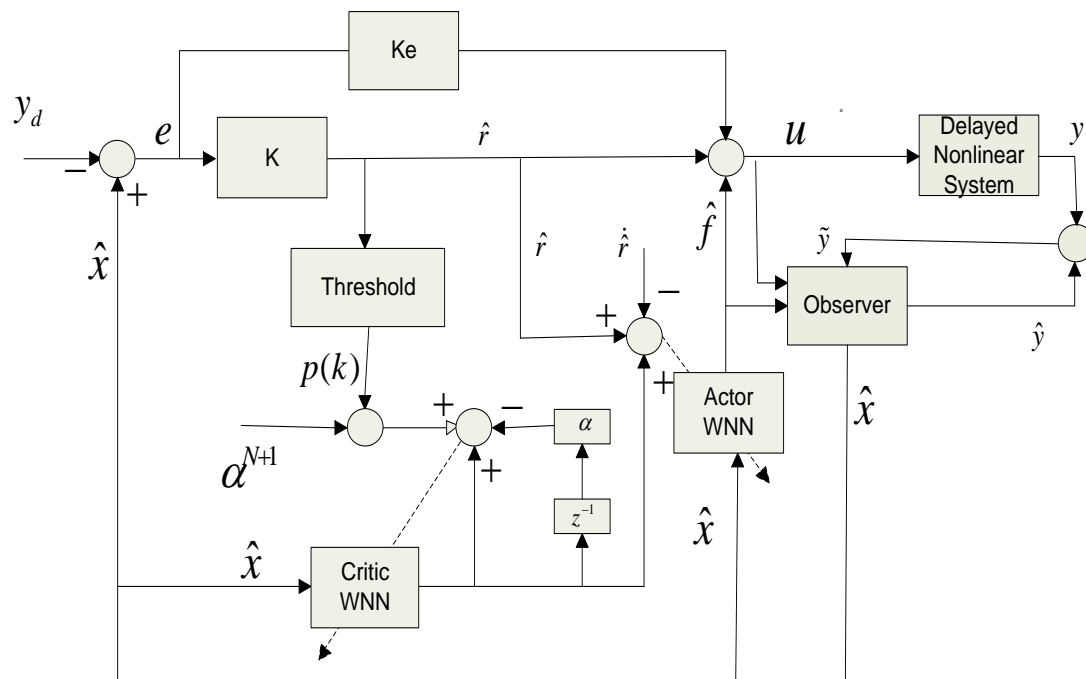


Figure 4. Overall closed loop system using Actor Critic WNN

How to cite this paper: Manish Sharma, Ajay Verma, "Wavelet Neural Network Observer Based Adaptive Tracking Control for a Class of Uncertain Nonlinear Delayed Systems Using Reinforcement Learning", International Journal of Intelligent Systems and Applications (IJISA), vol.4, no.2, pp.28-34, 2012. DOI: 10.5815/ijisa.2012.02.03