# ROBUST ADAPTIVE CONTROL USING REINFORCEMENT LEARNING FOR NONLINEAR SYSTEM WITH INPUT CONSTRAINTS

**Nguyen Tan Luy**[(1)], **Nguyen Thien Thanh**[(1)], **Nguyen Thi Phuong Ha**[(2)]
(1) National Key Lab of Digital Control and System Engineering, VNU-HCM
(2) University of Technology, VNU-HCM

***ABSTRACT:*** *This paper proposes a novel approach to design a controller in discrete time for the class of uncertain nonlinear systems in the presence of magnitude constrains of control signal which are treated as the saturation nonlinearity. A associative law between reinforcement learning algorithm based on adaptive NRBF neural networks and the theory of robust control $H_\infty$ is set up in a novel control structure, in which the proposed controller allows learning and control on-line to compensate multiple uncertain nonlinearities as well as minimizing both the $H_\infty$ tracking performance index function and the unknown nonlinear dynamic approximation errors. The novel theorem of robust stabilization of the closed-loop system is declared and proved. Simulation results verify the theoretical analysis.*

*Keywords: Reinforcement learning, robust control, neural network control*

## 1. INTRODUCTION

Direct adaptive controllers for a class of nonaffine and affine uncertain nonlinear discrete-time systems with input constraints using reinforcement learning neural networks are proposed in [1-2]. The performance index functions of the long term tracking error are predicted and minimized by reinforcement learning algorithms. As results, some of nonlinear components such as unknown dynamic functions, the control inputs constrained saturation and unknown but bounded disturbance are compensated. In addition, the tracking error and the functional approximation error of neural networks are uniform ultimate bounded (UUB) using Lyapunov approach.

In the theory of robust control, available knowledge of system is to exploit absolutely such as nominal models or the upper bound of uncertain parameters to design robust stable controllers. However the robust controllers trend to become "hard" controller because they contain constant parameters. On the other hand, reinforcement learning (RL) methods can be learn online to find better control laws without the available knowledge. However, RL methods deal with processes of try and error, therefore at the intermediate stage of learning and control the RL systems may go through periods of unstable behavior.

Recently, to solve the above problem, some methods of robust RL have been proposed as (1) a RL algorithm using neural networks (NN) combines with the concept of sliding mode control [4]. This method makes the system be oscillated by the chattering phenomenon, although the learning system is robust. (2) A tool of robust control theory, Integral Quadratic Constraints (IQCs), is used in robust reinforcement learning [5-6]. By replacing the nonlinear and time-varying components of the NNs with IQCs, NN's weights are analyzed and constrained in stable dynamic ranges. As results, NNs generate control signals which make the system be robust stable during learning and control online. (3) Another method is designed based on theory of $H_\infty$ control for the system whose modeling errors can be pre-interpreted as unknown but bounded disturbance [7]. The main purpose of this method is that an online

function known as *Hamilton-Jacobi-Isacc (HJI)* is approximated to drive the worst disturbance and the optimal control simultaneously.

This paper contributes some novel points of view as follows

Combining a reinforcement learning algorithm based on neural networks and the theory of $H_\infty$ control to propose a novel robust adaptive control structure diagram for a class of the nonlinear discrete time system with input constrains.

The new robust adaptive reinforcement learning controller is analyzed and designed.

The new robust stable theorem is shown and proved.

The remainder of this paper is arranged as follows. Section 2 describes properties of the function approximator using NN as adaptive normalized RBF. A description of the uncertain nonlinear discrete time system with input constrains is presented in section 3. Small gain theorem in robust control theory is reviewed in section 4. In section 5, a novel control structure diagram is shown and a novel theorem of robust stabilization of the closed-loop system is declared and proved, subsequently. The results of simulation in section 6 verify the effects of the proposed controller and conclusions are drawn in section 7.

## 2. APPROXIMATION PROPERTY OF ADAPTIVE NORMALIZED RBF -ANRBF

Choosing suitable function approximators in RL is essential for speeding up learning and control. *ANRBF* with ability to adapt centers and widths of basic functions give approximation performance better than other neural networks [8].

A continuous function $f(x(k)) \in C(S)$ within a compact subset $S \in \Re^n$ is approximated by *ANRBF* as

$$f(x(k)) = W^T \Phi(x(k)) + \varepsilon(x(k)) \tag{1}$$

Where $W$ is a target weight matrix of the hidden layer to the output; $\Phi(x(k))$ is vector of the basis functions at instant $k$; $\varepsilon(x(k))$ is vector of functional approximation error. The actual *ANRBF* output is defined as

$$\hat{f}(x(k)) = \hat{W}^T(k) \Phi(x(k)) \tag{2}$$

$\hat{W}(k)$ is a weight matrix updated online at instant $k$; $n_h$ is number of hidden-layer nodes, an element $j^{th}$ of $\Phi(x(k))$ is defined as

$$\Phi_j(x(k)) = \frac{e^{-\frac{\|c_j - x(k)\|^2}{\sigma_j^2}}}{\sum_{l \in n_h} e^{-\frac{\|c_l - x(k)\|^2}{\sigma_l^2}}}, \quad j = 1, 2, \ldots, n_h$$

Where $c_j \in \Re^{n_i}$, $\sigma_j \in \Re$ denotes the vector of center and the value of width of $\Phi_j(x(k))$ respectively and $n_i$ is number of input-layer nodes.

*Remark:* with limited $n_h$, the following inequality is always satisfied

$$\alpha \Phi^T(x(k)) \Phi(x(k)) \leq 1, \forall \alpha : \alpha \leq \frac{1}{n_h} \tag{3}$$

## 3. UNCERTAIN NONLINEAR DISCRETE TIME SYSTEM DESCRIPTION

Consider the following uncertain nonlinear discrete time system

$$x_1(k+1) = x_2(k)$$
$$\vdots \tag{4}$$
$$x_n(k+1) = f(x(k)) + u(k) + d(k)$$

Where $x(k) = \left[ x_1^T(k), x_2^T(k), ..., x_n^T(k) \right]^T \in \Re^{nm}$, $x_i(k) \in \Re^m$, $i = 1, ..., n$ is the vector of state at instant $k$; $f(x(k)) \in \Re^m$ is the unknown nonlinear dynamics of the system; $u(k) \in \Re^m$ is the control input constrained saturation; and $d(k) \in \Re^m$ is the unknown but bounded disturbance.

Given a reference trajectory $x_{nd}(k) \in \Re^m$ and its past values, the vector of tracking error $e_i(k) \in \Re^m$ is defined as

$$e_i(k) = x_i(k) - x_{nd}(k+i-n) \tag{5}$$

Define the filtered tracking error $r(k) \in \Re^m$ as

$$r(k) = [\Lambda\ I] e(k) \tag{6}$$

$e(k) = \left[ e_1^T(k), e_2^T(k), ..., e_n^T(k) \right]^T$; $e_1(k) = e_2(k)$ With $e_1(k+1)$ is the next value of $e_1(k)$; $e_{n-1}(k), ..., e_1(k)$ are the past values for $e_n(k)$; $I \in \Re^{m \times m}$ is an identity matrix; $\Lambda = [\lambda_{n-1}, \lambda_{n-2}, ..., \lambda_1] \in \Re^{m \times (n-1)m}$ with $\lambda_i \in \Re^{m \times m}$, $i = 1, ..., (n-1)$ is constant diagonal positive definite matrix chosen so that its eigenvalues are within the unit circle. Consequently, if $\lim\limits_{k \to \infty} r(k) = 0$ then $e(k)$ will go to zero. Combining (4) (5) and (6) we get

$$r(k+1) = f(x(k)) - x_{nd}(k+1) + \lambda_1 e_n(k) + ..., + \lambda_{n+1} e_2(k) + u(k) + d(k) \tag{7}$$

The control purpose is to make the tracking error of the system (7) achieve the $H_\infty$ robust performance index.

## 4. $H_\infty$ CONTROL FOR DISCRETE TIME SYSTEMS

$H_\infty$ Robust control deals with a system shown in Fig. 1, where $G$ is the controlled plant, $K$ is the controller, $u(k)$ is the control input, $y(k)$ is the output of plant supposed measurement available to the controller. The controller $K$ is designed to stabilize the closed loop system based on model $G$. However there is difference between the model and actual plant dynamics, the feedback loop could be unstable. The effect of modeling error can be seen as an unknown disturbance $\xi(k) \in L_2[0, \infty]$ generated by unknown mapping $\Delta$ from $e(k)$ to $\xi(k)$.

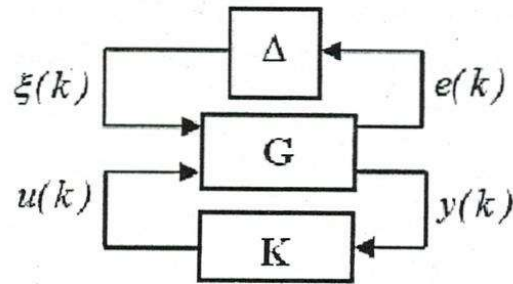According to the Small Gain Theorem, the system in Fig. 1 will be stable if the condition as follows is satisfied.

**Figure 1.** The Small Gain Theorem

$$\sum_{k=0}^{N} \|e(k)\|^2 \le \rho^2 \sum_{k=0}^{N} \|\xi(k)\|^2 + \eta \tag{8}$$

Where $\|\Delta\|_\infty < \dfrac{1}{\rho}$, $\rho$ is a specified attenuation level; $\eta$ is the positive constant depending on initial conditions; $N$ is number of steps to the final state.

## 5. DESIGN ROBUST ADAPTIVE REINFORCEMENT LEARNING CONTROLLER

### 5.1. Basic control law

At early stages of learning online, the control loop using NN whose weights are selected random from $[0,1]$ will be unstable. Therefore using a basic control law to make system be stable is necessary [1-2]. This control law provides the supervised signals which allow the reinforcement learning system turning NN's weights online rather than offline training. To find it, the auxiliary control input $v(k) \in \Re^m$ is defined as

$$v(k) = x_{nd}(k+1) - \hat{f}(x(k)) - \lambda_1 e_n(k) - \ldots, -\lambda_{n+1} e_2(k) + Lr(k) \tag{9}$$

Where $\hat{f}(x(k))$ is the function approximation of $f(x(k))$ and $L \in \Re^{m \times m}$ is a diagonal matrix. The actual control input constrained saturation is defined

$$u(k) = \begin{cases} v(k), & \|v(k)\| \le u_{max} \\ u_{max}\, sgn(v(k)), & \|v(k)\| > u_{max} \end{cases} \tag{10}$$

Where $u_{max} \in \Re$ is the upper bound for $u(k)$. The closed loop system can be written as

$$r(k+1) = Lr(k) - \tilde{f}(x(k)) + d(k) + \Delta u(k) \tag{11}$$

And

$$\tilde{f}(x(k)) = \hat{f}(x(k)) - f(x(k)) \tag{12}$$

$\Delta u(k) = u(k) - v(k)$ , $\Delta u(k) \in \Re^m$ . Two auxiliary signals $\overline{e}(k) \in \Re^m$ and $e_u(k) \in \Re^m$ are defined as follows to reject the effect of $\Delta u(k)$

$$\overline{e}(k+1) = L\overline{e}(k) + \Delta u(k) \tag{13}$$

$$e_u(k) = r(k) - \overline{e}(k) \tag{14}$$

Combining (11) (13) and (14) we get

$$e_u(k+1) = Le_u(k) - \tilde{f}(x(k)) + d(k) \tag{15}$$

Where $Le_u(k)$ is the basic control law.

## 5.2. Robust Adaptive Reinforcement Learning -RARL

Fig. 2 represents a RARL system based on the special structure known as actor-*critic* [2-3]. Here, the *actor* and *critic* are based on *ANRBFs*.

*Remark: $e(k)$* in Eq. (8) is replaced by $e_u(k)$ and $\xi(k)$ is defined as

$$\xi(k) = \varepsilon(x(k)) + d(k) \tag{16}$$

Where $\varepsilon(x(k))$ is the total of the functional approximation error of both *actor* and *critic*.

### 5.2.1. Value function

The performance index $J(k) \in \Re^m$ at instant $k$ is proposed as

$$J(k) = \begin{cases} 0, & e_u{}^T(k)e_u(k) - \rho^2 d^T(k)d(k) \leq 0 \\ 1, & e_u{}^T(k)e_u(k) - \rho^2 d^T(k)d(k) > 0 \end{cases} \tag{17}$$

And the value function at instant $k$ becomes

$$Q(k) = \sum_{i=k}^{N} \gamma^{N-i+k} J(i) \tag{18}$$

Where $\gamma, (0 < \gamma < 1)$, is a discount factor which makes $Q(k)$ converge when $N \to \infty$. The optimal value function $Q^*(k)$ satisfies the Optimal Bellman Equal as

$$Q^*(k) = \min_{u(k)} \left( \gamma Q^*(k-1) - \gamma^{N+1} J(k) \right) \tag{19}$$

Solution of Eq. (19) could not be found by analytic or the bellman meshed diagram because the model is not available. Hereafter $Q^*(k)$ is approximated based on the *actor-critic* system, in which the output of *critic* is used to approximate $Q(k)$, and the output of *actor* generates the control signal to approximate $Q^*(k)$. Weights of *actor* are updated by the signal from *critic*.

### 5.2.2. Critic

The *critic* is used to approximate the value function $Q(k)$ to $\hat{Q}(k)$. In RL, the prediction error [10] is defined as

$$\varepsilon_c(k) = \alpha^{N+1} J(k) - \alpha \hat{Q}(k-1) + \hat{Q}(k) \tag{20}$$

$$\hat{Q}(k) = \hat{W}_c(k)\Phi_c(x(k)) \tag{21}$$

And $\varepsilon_c(x(k)) \in \Re^m$; $\hat{Q}(k) \in \Re^{nm \times n_c}$; $\hat{W}_c(k) \in \Re^{n_c \times m}$ is the weight matrix, $\Phi_c(k) \in \Re^{n_c}$ is the vector of actor functions, $n_c$ is the number of hidden-layer nodes, $x(k) \in \Re^{nm}$ is the input to the *critic*. The law for updating weights is proposed as

$$\begin{aligned} \Delta\hat{W}_c(k) = -\alpha_c \Phi_c(x(k))\varepsilon_c(k) = &-\alpha_c \Phi_c(x(k))\big(\hat{W}_c(k)\Phi_c(x(k)) \\ &+ \alpha^{N+1} J(k) - \hat{W}_c(k+1)\Phi_c(x(k-1))\big)^T \end{aligned} \tag{22}$$

Where $\alpha_c \in \Re$ is the positive constant representing learning rate.

### 5.2.3. Actor

The function $f(x(k))$ in Eq. (4) is approximate to $\hat{f}(x(k))$ by the *Actor*. It can be seen $\hat{f}(x(k))$ as an optimal control input $u^*(k)$ which makes $Q(k)$ converge to $Q^*(k)$

$$\hat{f}(x(k)) = \hat{W}_a^T(k)\Phi_a(x(k)) \tag{23}$$

Where $\hat{W}_a(k) \in \Re^{n_a \times m}$, $\Phi_a(k) \in \Re^{n_a}$, $n_a$, $x(k) \in \Re^{nm}$ are the weight matrix, the vector of actor functions, number of hidden-layer nodes and the input to the *actor* respectively. The law updating weights is proposed as

$$\Delta\hat{W}_a(k+1) = -\alpha_a \Phi_a(x(k)) \times \left(\hat{Q}(k) + Le_u(k) - e_u(k+1)\right) \tag{24}$$

Where $\alpha_a \in \Re$ is the positive constant representing learning rate.

### 5.2.4. Robust stability

*Theorem*: given the bounded reference trajectory $x_{nd}(k)$ and its past value, defined the auxiliary control input in Eq. (9), the $L_{max}$ is the maximum singular value of the gain matrix $L$ in Eq. (15) satisfies as

$$L_{max} < \frac{\sqrt{3}}{3} \tag{25}$$

And the function of performance index in Eq. (17), *actor-critic* structure base on *ANRBF*, the laws of updating weight for *critic* as Eq. (22) and *actor* as Eq. (24) then during learning and control online, the tracking error of the closed loop system will be achieving the $H_\infty$ robust stability.

*Proof*: See the Appendix.

## 6. SIMULATION

Nonlinear system for simulation to verify proposed controller is given by Eq. (26)

$$\begin{aligned} x_1(k+1) &= x_2(k) \\ x_2(k+1) &= f(x(k)) + u(k) + d(k) \end{aligned} \tag{26}$$

Where $f(x(k)) = -\dfrac{5}{8}\left(\dfrac{p_1 x_1(k)}{1 + p_2 x_2^{p_3}}\right) + 0.3x_2(k)$

$p_i \in \Re$ $i = 1,2,3$ is uncertain parameter bounded as $p_1 = [-5,5]$, $p_2 = [-1,1]$, $p_3 = [0,5]$ respectively. The control objective is to design *RARL* so that $x_2$ tracks desired trajectory $x_{2d}$ with considering saturated gain phenomenon of the control input. $x_{2d}$ is given as

$$x_{2d} = \begin{cases} \sin\left(0.1kT + \dfrac{\pi}{2}\right), & 0 \le k \le 3000 \\ -1, & 3000 < k \le 4000 \\ -1, & 5000 < k \le 6000 \\ 1, & 4000 < k \le 5000 \end{cases} \tag{27}$$
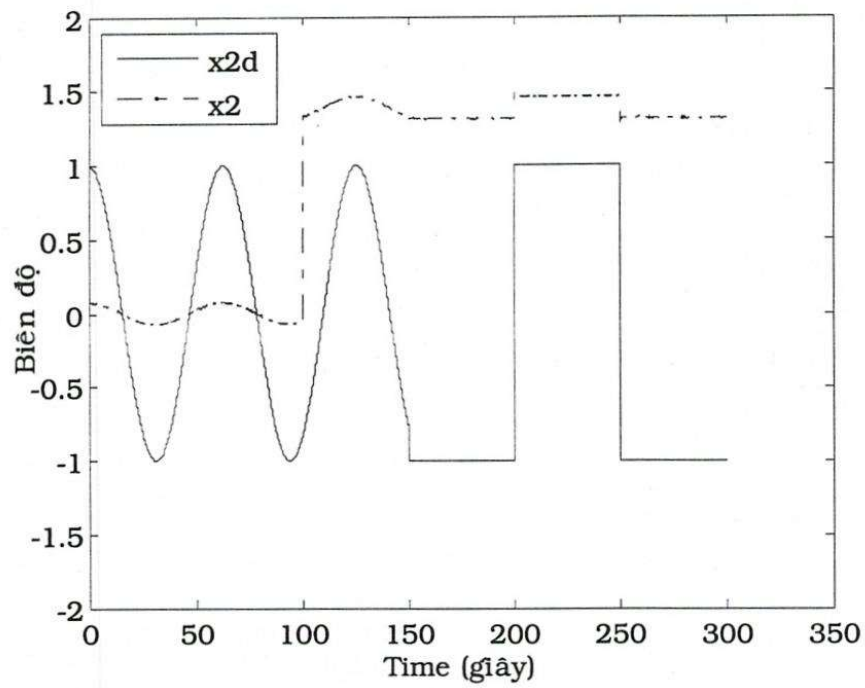
**Fig. 3.** Performance of the basic controller

$$( p_1 = 1, p_2 = 1, p_3 = 2 )$$
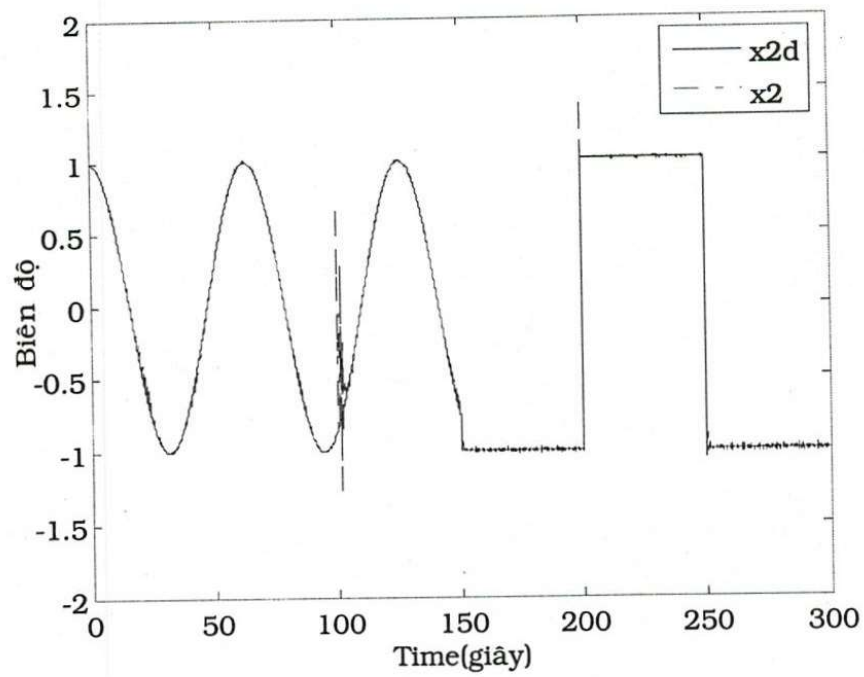


**Fig. 4.** Performance of *RARL*

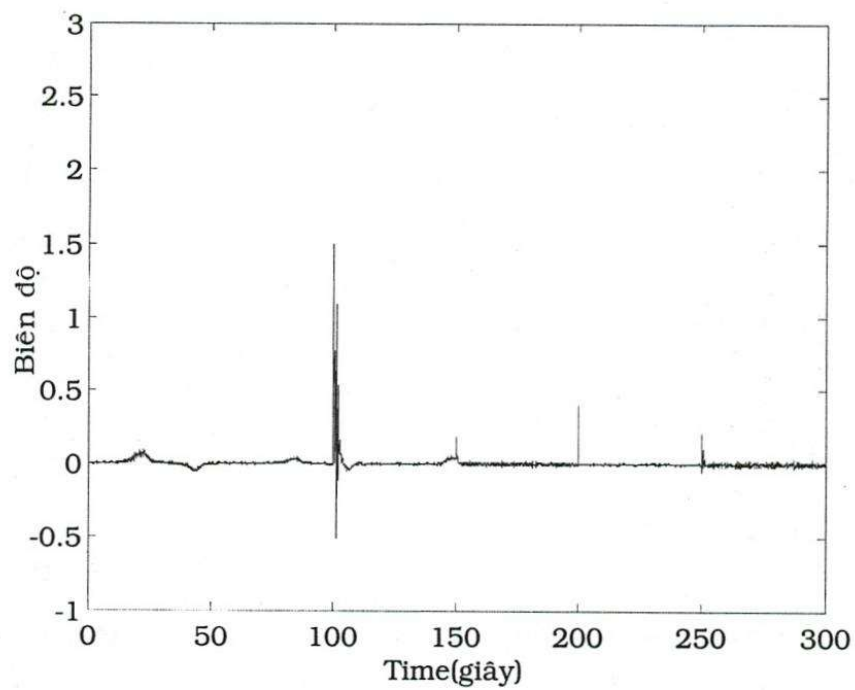$$( p_1 = 1, p_2 = 1, p_3 = 2 )$$

**Fig. 5.** Tracking error of *RARL* ( $p_1 = 1, p_2 = 1, p_3 = 2$ )
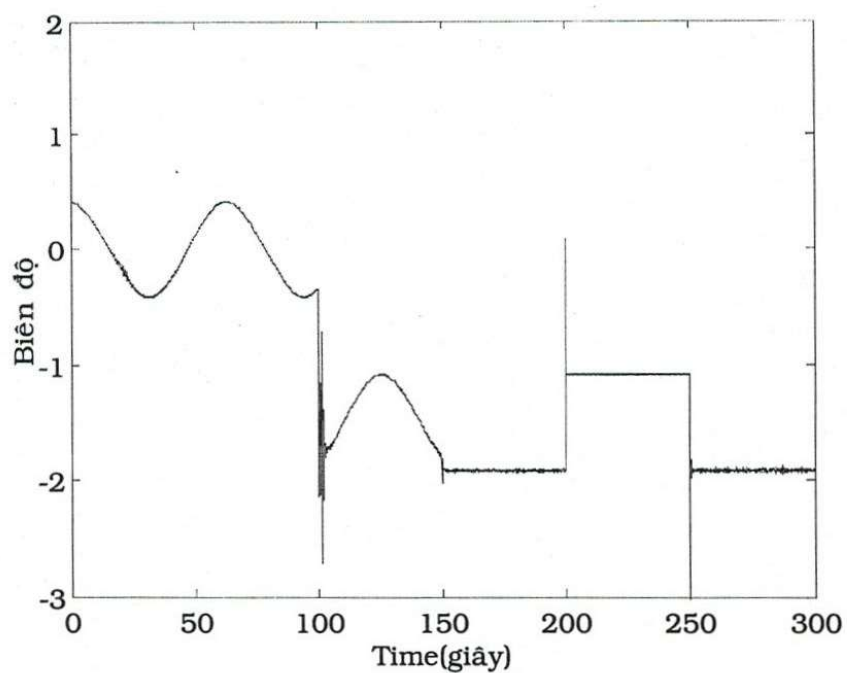


**Fig. 6.** Control input of *RARL*
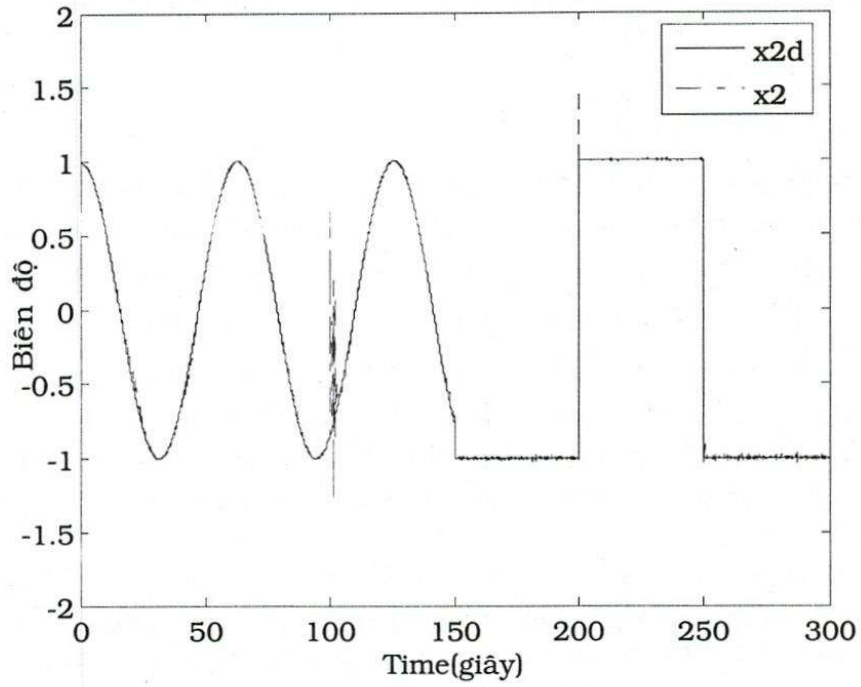
( $p_1 = 1, p_2 = 1, p_3 = 2$ )
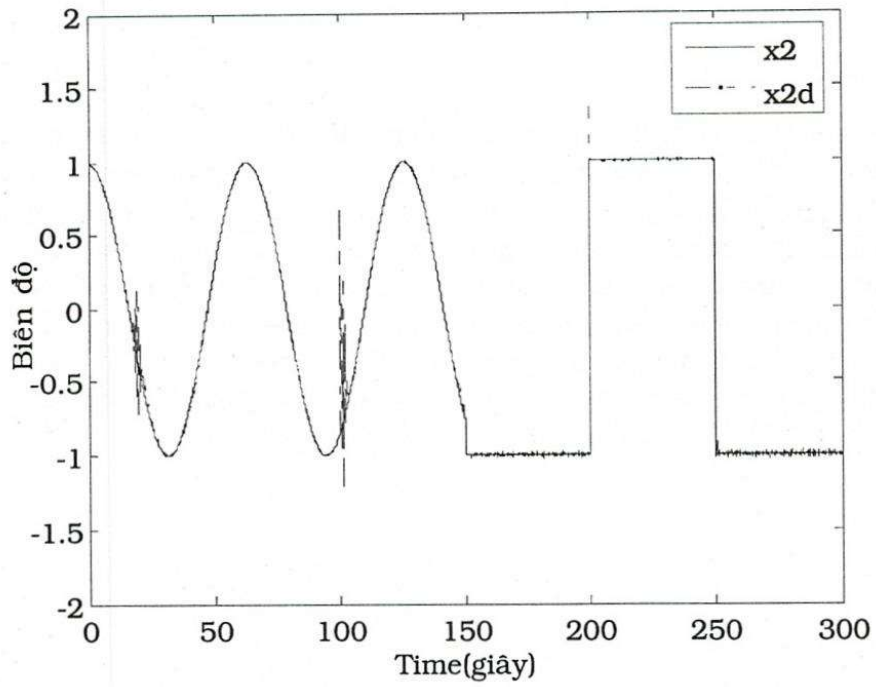
**Fig. 7.** Performance of RARL



**Fig. 8.** Performance of *RARL*

$$( p_1 = -1, p_2 = 1, p_3 = 5 )$$

The sampling interval is taken as $T = 0.05s$ and the white Gaussian noise with a standard deviation of 0.005 is added to the system. The time duration of simulation is taken 300s. The unknown disturbance is chosen as

$$d(k) = \begin{cases} 0 & k < 2000 \\ 1.5, & 2000 \le k \le 6000 \end{cases}$$

The gain of the basis control input is chosen as $L = -0.15$. Both number of hidden-layer nodes of *critic* and *actor* are selected as $n_c = n_a = 10$; the update rates are $\alpha_c = \alpha_a = 0.1$. The activation functions are the same for both of them, where $\sigma_i = 0.1$, $i = 1,...,10$, $x_{mi}$ is uniformly partitioned within $[-1,1]$. All of weights are initialized at random from $[0,1]$. The gain of the control input is constrained within $[-3,3]$. The discount factor is selected as $\gamma = 0.5$.

First, to show the effect of proposed controller, the *RARL* controller is removed out of the closed loop. The uncertainty parameters are selected as $p_1 = 1, p_2 = 1, p_3 = 2$. In Fig.3 it can be seen that the tracking error given by the basic controller are bounded but the performance is very poor.

Now we add the RARL controller to the closed loop. The $H_\infty$ robust tracking performance is presented in Fig. 4 and Fig. 5. Because of the activation of disturbance, at the second of 100 the tracking error $e(k)$ is overshoot but it quickly goes to zero asymptotically. The Fig. 6 presents the control input in which its gain is constrained in range of $\pm 3$.

The $H_\infty$ robust tracking performance with $p_1 = -5, p_2 = -1, p_3 = 0$ is presented in Fig. 7, with $p_1 = 1, p_2 = 1, p_3 = 2$ in Fig. 8. From Fig. 4, 7 and 8 we can see that they are robust for all the parameters.

## 7. CONCLUSION

This paper proposes the method which combined reinforcement learning based on neural network *ANRBF* and the robust theory $H_\infty$ to design a robust adaptive reinforcement learning controller for a class of the uncertain nonlinear discrete time system with input constrains which are treated as the saturation nonlinearity. The proposed controller not only compensates some uncertainty nonlinear components but also gives the robust tracking performance.

An Adaptive controller with robust tracking performance using the recurrent *CMAC* neural network to get rid of chattering phenomenon for a class of multivariable uncertain nonlinear system is proposed in [9]. Develop a RARL controller using *CMAC* is next research. In addition applying RARL to control for real plants is considered next.

# ĐIỀU KHIỂN THÍCH NGHI BỀN VỮNG DÙNG HỌC CỦNG CỐ CHO HỆ THỐNG PHI TUYẾN VỚI RÀNG BUỘC NGÕ VÀO

**Nguyễn Tấn Lũy**[(1)]**, Nguyễn Thiện Thành**[(1)]**, Nguyễn Thị Phương Hà**[(2)]
(1) PTN Trọng điểm Quốc gia Điều khiển số & Kỹ thuật hệ thống, ĐHQG-HCM
(2) Trường Đại học Bách Khoa, ĐHQG-HCM

***TÓM TẮT:*** *Bài báo đề xuất phương pháp mới để thiết kế bộ điều khiển thích nghi bền vững cho lớp hệ thống phi tuyến rời rạc bất định với ràng buộc về biên độ của tín hiệu điều khiển được xử lý như là độ phi tuyến bão hòa. Luật kết hợp giữa thuật toán học củng cố sử dụng mạng thần kinh nhân tạo NRBF thích nghi và lý thuyết khiển bền vững $H_\infty$ được thiết lập trong cấu trúc điều khiển mới trong đó bộ điều khiển đề xuất cho phép học và điều khiển trực tuyến để bù đa thành phần phi tuyến cũng như tối thiểu phiếm hàm chỉ tiêu chất lượng bám $H_\infty$ và sai số ước lượng động phi tuyến không biết. Định lý mới về sự ổn định bền vững của hệ thống vòng kín được phát biểu và chứng minh. Kết quả mô phỏng đã kiểm chứng các phân tích về lý thuyết.*

***Từ khóa****: Reinforcement learning, robust control, neural network control*

## REFERENCES

[1]. Yang Q. and Jagannathan S., Online Reinforcement Learning Control of Unknown Nonaffine Nonlinear Discrete Time Systems, *IEEE Conf. Decision and Control*, 2007.

[2]. He P. and Jagannathan S., Reinforcement learning neural-network-based controller for nonlinear discrete-time systems with input constraints, *IEEE Trans. System, Man., and Cybernetic*, 2007.

[3]. Lin F., *Robust Control Design: An Optimal Control Approach*, John Wiley & Sons, Ltd, 2007.

[4]. Obayashi M., Nakahara N. Kuremoto T., Kobayashi K., A robust reinforcement learning using the concept of sliding mode control, *13th International Symposium on Artificial Life and Robotics*, 2009.

[5]. Anderson C. W. and Young P. M., Robust reinforcement learning control using integral quadratic constraints for recurrent neural networks, *IEEE Trans. Neural Netw.*, 2007.

[6]. Kretchmar R. M., *A synthesis of reinforcement learning and robust control theory*, Ph.D. dissertation, Dept. Comput. Sci., Colorado State Univ., Fort Collins, CO, 2000.

[7]. Morimoto J. and Doya K., Robust reinforcement learning, *Neural Comput.*, 2005.

[8]. Ha N.T.P., Thanh N.T., Luy N.T., Research on function aproximators in supervised and reinforcement learning, *journal of science and technology*, 2008

[9]. Lin C.M, Lin M.H and Chen C.H., $H_\infty$-Tracking based adaptive control for a class of multivariable uncertain nolinear system using recurrent CMAC, *IAENG International Journal of Computer Science*, 2007.

[10]. J. Si and Y. T. Wang, On-line learning control by association and reinforcement, *IEEE Trans. Neural Netw.*, vol. 12, no. 2, pp. 264–276, 2001

## APPENDIX

Proof of the theorem in section 5.2.4:

Remark: in following equations $x(k)$ is replaced by $k$

Selected the Lyapunov function candidate

$$V(k) = \frac{1}{\gamma_1} e_u^T(k) e_u(k) + \frac{1}{\alpha_c} tr\left(\tilde{W}_c^T(k)\tilde{W}_c(k)\right) + \frac{1}{\gamma_2 \alpha_a} tr\left(\tilde{W}_a^T(k)\tilde{W}_a(k)\right)$$
$$+ \frac{1}{\gamma_3}\left\|\psi_c(k-1)\right\|^2 + \frac{4}{\gamma_2}(N+1-k)W_{cmax}^2 \Phi_{cmax}^2 \tag{28}$$

Where $W_{cmax}$, $\Phi_{cmax}$ denote the upper bounds of $W_c$ and $\Phi_c$ respectively. $\gamma_i > 0, i = 1,2,3$; $\left. \psi_c(k-1)\right|_{k=0} = 0$, N is the final step.

The differences of $\Delta V(k)$ form Eq. (28) is decomposed as

$$\Delta V(k) = \Delta V_1(k) + \Delta V_2(k) + \Delta V_3(k) + \Delta V_4(k) + \Delta V_5(k) \tag{29}$$

Combining (1),(2)(15) and (23) we get

$$\Delta V_1(k) = \frac{1}{\gamma_1}\left(\left(Le_u(k) - \psi_a(k) + \varepsilon_a(k) + d(k)\right)^T\right.$$
$$\times \left(Le_u(k) - \psi_a(k) + \varepsilon_a(k) + d(k)\right) - e_u^T(k)e_u(k)\right) \tag{30}$$
$$\leq \frac{3}{\gamma_1}\left((L_{max}^2(L) - \frac{1}{3})\left\|e_u(k)\right\|^2 + \left\|\psi_a(k)\right\|^2 + \left\|\varepsilon_a(k) + d(k)\right\|^2\right)$$

Where $\psi_a(k) = \tilde{W}_a(k)\Phi_a(k)$; $L_{max} \in \Re$ is the maximum singular value of $L$.

$\Delta V_2(k)$ is presented as

$$\Delta V_2(k) = \frac{1}{\alpha_c} tr\left(\tilde{W}_c^T(k+1)W_c(k+1) - \tilde{W}_c^T(k)\tilde{W}_c(k)\right) \tag{31}$$

Substituting (22) into (31) and rewrite them we get

$$\Delta V_2(k) \leq -\left(1 - \alpha_c \Phi_c^T(k)\Phi_c(k)\right)$$
$$\left\|\psi_c(k) + W_c^T(k)\Phi_c(k) + \gamma^{N+1}J(k) - \gamma\hat{W}_c^T(k-1)\Phi_c(k-1)\right\|^2$$
$$- \left\|\psi_c(k)\right\|^2 + 2\left\|W_c^T(k)\Phi_c(k) + \gamma^{N+1}J(k) - \gamma W_c^T(k-1)\Phi_c(k-1)\right\|^2 \tag{32}$$
$$+ 2\gamma^2\left\|\psi_c(k-1)\right\|^2$$

$\Delta V_3(k)$ is presented as

$$\Delta V_3(k) = \frac{1}{\gamma_2} tr\left(\tilde{W}_a^T(k+1)W_a(k+1) - \tilde{W}_a^T(k)\tilde{W}_a(k)\right) \tag{33}$$

Substituting (24) into (33) and rewrite them we get

$$\Delta V_3(k) \le \frac{1}{\gamma_2}\left\{-\left(1-\alpha_a\Phi_a^T(k)\Phi_a(k)\right)\left\|\psi_a(k)+\hat{W}_c^T(k)\Phi_c(k)\right.\right.$$

$$\left.-(\varepsilon_a(k)+d(k))\right\|^2 - \left\|\psi_a(k)\right\|^2\right\} \tag{34}$$

$$+\frac{2}{\gamma_2}\left\|W_c^T(k)\Phi_c(k)-(\varepsilon_a(k)+d(k))\right\|^2+\left\|\psi_c(k)\right\|^2\right\}$$

$\Delta V_4(k)$ and $\Delta V_5(k)$ are presented as

$$\Delta V_4 = \frac{1}{\gamma_3}\left(\left\|\psi_c(k)\right\|^2 - \left\|\psi_c(k-1)\right\|^2\right) \tag{35}$$

$$\Delta V_5 = \frac{4}{\gamma_2}(N+1-(k+1))W_{cmax}^2\Phi_{cmax}^2 - \frac{4}{\gamma_2}(N+1-k)W_{cmax}^2\Phi_{cmax}^2 = -\frac{4}{\gamma_2}W_{cmax}^2\Phi_{cmax}^2 \tag{36}$$

Now, substituting (30), (32), (34), (35) and (36) into (29) we get

$$\Delta V(k) \le -\frac{1}{\gamma_1}(1-3L_{max}^2)\left\|e_u(k)\right\|^2 - \left(1-\frac{1}{\gamma_3}-\frac{2}{\gamma_2}\right)\left\|\psi_c(k)\right\|^2 - \left(\frac{1}{\gamma_2}-\frac{3}{\gamma_1}\right)\left\|\psi_a(k)\right\|^2$$

$$-\left(\frac{1}{\gamma_3}-2\gamma^2\right)\left\|\psi_c(k-1)\right\|^2$$

$$-\left(1-\alpha_c\Phi_c^T(k)\Phi_c(k)\right)\left\|\psi_c(k)+W_c^T(k)\Phi_c(k)+\gamma^{N+1}J(k)\right.$$

$$\left.-\gamma\hat{W}_c^T(k-1)\Phi_c(k-1)\right\|^2 - \frac{1}{\gamma_2}\left\{\left(1-\alpha_a\Phi_a^T(k)\Phi_a(k)\right)\right.$$

$$\times\left\|\psi_a(k)+\hat{W}_c^T(k)\Phi_c(k)-\left(\varepsilon_a(k)+d(k)\right)\right\|^2\right\}$$

$$+2\left\|W_c^T(k)\Phi_c(k)+\gamma^{N+1}J(k)-\gamma W_c^T(k-1)\Phi_c(k-1)\right\|^2$$

$$+\frac{2}{\gamma_2}\left\|W_c^T(k)\Phi_c(k)-\left(\varepsilon_a(k)+d(k)\right)\right\|^2+\frac{3}{\gamma_1}\left\|\varepsilon_a(k)+d(k)\right\|^2-\frac{4}{\gamma_2}W_{cmax}^2\Phi_{cmax}^2 \tag{37}$$

Selecting $\gamma_1 > 3\gamma_2$, $\gamma_2 > \dfrac{2}{1-2\gamma^2}$, $\gamma_3 = \dfrac{1}{2\gamma^2}$, $L_{max} < \dfrac{\sqrt{3}}{3}$ and simplifying we get

$$\Delta V(k) \le -\frac{1}{\gamma_1}(1-3L_{max}^2)\|e_u(k)\|^2$$

$$+2\|W_c^T(k)\Phi_c(k)+\gamma^{N+1}J(k)-\gamma W_c^T(k-1)\Phi_c(k-1)\|^2$$

$$+\frac{2}{\gamma_2}\|W_c^T(k)\Phi_c(k)-(\varepsilon_a(k)+d(k))\|^2$$

$$+\frac{3}{\gamma_1}\|\varepsilon_a(k)+d(k)\|^2-\frac{4}{\gamma_2}W_{cmax}^2\Phi_{cmax}^2$$

(38)

Combining (20)(21) and (22) to (38) we get

$$\Delta V(k) \le -\frac{1}{\gamma_1}(1-3L_{max}^2)\|e_u(k)\|^2+2\|\varepsilon_c(k)\|^2+\frac{4}{\gamma_2}\|W_c^T(k)\Phi_c(k)\|^2$$

$$+\left(\frac{3}{\gamma_1}+\frac{4}{\gamma_2}\right)\|\varepsilon_a(k)+d(k)\|^2-\frac{4}{\gamma_2}W_{cmax}^2\Phi_{cmax}^2$$

(39)

$$\le -\frac{1}{\gamma_1}(1-3L_{max}^2)\|e_u(k)\|^2+\left(\sqrt{2}+\sqrt{\frac{3}{\gamma_1}+\frac{4}{\gamma_2}}\right)^2\|\varepsilon_c(k)+\varepsilon_a(k)+d(k)\|^2$$

Where $\dfrac{4}{\gamma_2}\|W_c^T(k)\Phi_c(k)\|^2 \le \dfrac{4}{\gamma_2}W_{cmax}^2\Phi_{cmax}^2$.

Taking sum of (39) we get

$$\sum_{k=0}^{N}\Delta V(k)=V(N+1)-V(0) \le -\frac{1}{\gamma_1}(1-3L_{max}^2)\sum_{k=0}^{N}\|e_u(k)\|^2$$

$$+\left(\sqrt{2}+\sqrt{\frac{3}{\gamma_1}+\frac{4}{\gamma_2}}\right)\sum_{k=0}^{N}\|\varepsilon_c(k)+\varepsilon_a(k)+d(k)\|^2$$

(40)

Apply (25) and remarking on $V(N+1)\ge 0$ and $V(0)\ge 0$ we get

$$\sum_{k=0}^{N}\|e_u(k)\|^2 \le \rho^2\sum_{k=0}^{N}\|\xi(k)\|^2+\eta$$

(41)

Where

$$\rho=\sqrt{\frac{\gamma_1}{1-3L_{max}^2}\left(\sqrt{2}+\sqrt{\frac{3}{\gamma_1}+\frac{4}{\gamma_2}}\right)}, \quad \|\xi(k)\|^2=\|\varepsilon_c(k)+\varepsilon_a(k)+d(k)\|^2, \quad \eta=\frac{\gamma_1}{1-3L_{max}^2}V(0)$$

The inequality (41) satisfies (8). So the proof of the theorem is given.