2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Optimal Control of Nonlinear Systems using RBF Neural Network and Adaptive Extended Kalman Filter

Peda V. Medagam, *Student Member, IEEE,* and Farzad Pourboghrat*, Senior Member, IEEE*

Abstract **— This paper presents a nonlinear optimal control technique based on approximating the solution to the Hamilton-Jacobi-Bellman (HJB) equation. The HJB solution (value function) is approximated as the output of a radial basis function neural network (RBFNN) with unknown parameters (weights, centers, and widths) whose inputs are the system's states. The problem of solving the HJB equation is therefore converted to estimating the parameters of the RBFNN. The RBFNN's parameters estimation is then recognized as an associated state estimation problem. An adaptive extended Kalman filter (AEKF) algorithm is developed for estimating the associated states (parameters) of the RBFNN. Numerical examples illustrate the merits of the proposed approach.**

I. Introduction

Optimal control is an important aspect of control theory because of guaranteed closed loop performance for a given system. Optimal control of linear time invariant (LTI) systems has been well studied and practiced with much success [1]. The optimal control design for LTI systems generally involves the solution of algebraic Riccati equation (ARE). One special extension to nonlinear systems involves state dependent Riccati equation (SDRE) technique, which provides high performance control. This method consists of changing the nonlinear system into a state dependent pseudo-linear form and involves the solution of an algebraic SDRE along the system's trajectory to obtain a nonlinear feedback controller [2].

For most nonlinear systems, however, the optimal control design requires the solution of Hamilton-Jacobi-Bellman (HJB) equation. The HJB equation in general cannot be solved analytically. There have been a number of approaches to solve the HJB equation. These approaches, generally, involve approximation techniques. One method of solution is based on power series approximation of HJB equation. The basic idea of this method is to approximate the value function as a truncated power series and to find the corresponding terms of the series by fitting it in the HJB equation [3], [4].

There are also other approximation techniques. Saridis *et al*. [5] developed a recursive approximation technique which starts with a stabilizing controller for a given plant and converges point-wise to the optimal control. Based on this technique Beard [6] proposed a successive Galerkin approximation for generalized Hamilton-Jacobi-Bellman (GHJB) equation and showed the convergence of the successive approximation for optimal control solution. The difficulty with this and other similar methods is the selection of basis functions, which are important for the convergence of the solution to optimal control. This difficulty may be resolved by employing wavelets [7] or neural networks [8] as basis functions.

Neural networks have been widely used in the identification, estimation and control of nonlinear systems [9], [10]. Offline estimation of the value function using neural networks has been studied in [11] where the neural network was trained using least square technique. In addition, nonlinear H_∞ control using radial basis function (RBF) neural networks has been reported in [12]. This method is based on the estimation of the value function using nonlinear RBF neural networks (RBFNN) where the network is trained, offline, using gradient method.

In the present paper, we solve the HJB equation for the optimal value function. The value function is needed to solve for an optimal control design. RBFNNs are used to estimate the value function as the solution to the generalized HJB (GHJB) equation. The performance of the RBFNNs depends on their weights, centers and widths. The common training approach is to first select the RBF centers using unsupervised K-means algorithm [13], or randomly from input data, and then determine the weights using least square or gradient descent method [12]. However, here, the weights as well as the centers and widths of the RBFNN are unknown and appear nonlinearly. An adaptive extended Kalman filter (AEKF) method is therefore developed to train the neural network weights, centers, and widths, online, with good accuracy. Fast digital signal processors (DSP) can be used for real-time implementation.

II. Statement of the problem

Consider a nonlinear time invariant system described by

$$\dot{x} = f(x) + g(x)u \tag{1}$$

where $x \in \Omega \subset \Re^n$, $u \in \Re^m$, $f : \Re^n \to \Re^n$, and $g : \Re^n \to \Re^n \times \Re^m$. Without loss of generality, it is assumed that $x_0 = 0$ is the equilibrium state and that $f(x_0) = 0$. Moreover the system is Lipschitz continuous on a set Ω

P. V. Medagam is with the Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, Carbondale, IL 62901-6603 USA (e-mail: pmedagam@phasetechnologies.com).

F. Pourboghrat is with the Department of Electrical and Computer Engineering, Southern Illinois University Carbondale, Carbondale, IL 62901-6603 USA (e-mail: pour@siu.edu).

978-1-4244-4524-0/09/$25.00 ©2009 AACC

in $\Re^n$. Also consider the cost function of the form

$$V(x) = \int_0^{\infty} [x^T Qx + u^T Ru]dt \qquad (2)$$

where $Q \in \Re^{n \times n}$ and $R \in \Re^{m \times m}$ are positive definite matrices. The optimal feedback control problem is to find the admissible control u^* so that the performance index (2) is minimized.

When an optimal control exists, it is given by

$$u^* = -\frac{1}{2} R^{-1} g(x)^T \nabla V^{*T}(x) \qquad (3)$$

where $\nabla = \frac{\partial}{\partial x}$ is the gradient operator, and that $V^*(x)$ is the value function and it satisfies the following Hamilton-Jacobi-Bellman (HJB) equation with boundary conditions $V^*(0) = 0$; i.e.,

$$v(\nabla V^*) = \nabla V^{*T} f + x^T Qx - \tfrac{1}{4} \nabla V^{*T} gR^{-1} g^T \nabla V^{*T} = 0 \qquad (4)$$

The above HJB equation is a two point boundary-value problem in terms of the value function $V^*(x)$, which in general is impossible to solve, analytically. The above equation can be also written in the form of generalized Hamilton-Jacobi-Bellman (GHJB) equation [3], i.e.,

$$v(\nabla V^*, u^*) = \nabla V^{*T} (f + gu^*) + x^T Qx + u^{*T} Ru^* = 0 \qquad (5)$$

which is a function of both ∇V^* and u^*. Unlike the HJB equation, the above GHJB equation is linear in terms of ∇V^{*T} and, together with equation (3), can be solved using approximation techniques.

III. Neural network approximation of the value function

Similar to the HJB equation, in general, the GHJB equation cannot be solved analytically. Hence, we approximate the value function with radial basis function neural networks. The RBF neural network has a feed-forward structure with one nonlinear hidden layer with N_z nodes and a linear output layer. Fig.1 shows the schematic diagram of the RBF neural network.

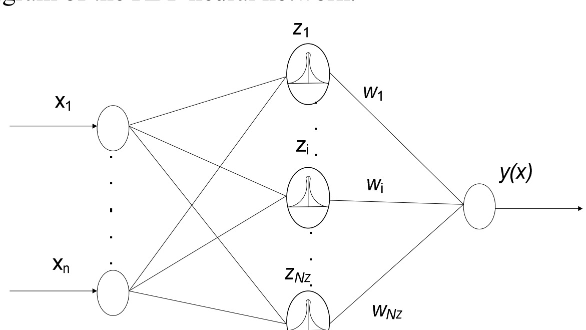

Fig. 1- Schematic diagram of the RBF neural network

The outputs of the hidden nodes are specified as

$$z_i = z\left(\|x - c_i\|; \sigma_i\right) \qquad (6)$$

where $i=1,..,N_z$, x is the input vector to the neural network, and that c_i's and σ_i's are, respectively, the adjustable centers and widths of the RBF functions. The activation functions, z_i's, are chosen as

$$z_i(x) = \left(\|x - c_i\|^2 + \sigma_i^2\right)^{-1/2} \qquad (7)$$

The output of the RBF neural network (RBFNN) is given by

$$y(x) = w^T z(x, c, \sigma) \qquad (8)$$

where $w^T = [w_1, \cdots, w_{N_z}]$, $c^T = [c_1, \cdots, c_{N_z}]$ and $\sigma^T = [\sigma_1, \cdots, \sigma_{N_z}]$ are the vectors of adjustable weights, centers, and widths of the RBFNN, and that $z^T(x) = [z_1, \cdots, z_{N_z}]$ is the vector of basis functions.

Realizing that the value function must be positive definite, we approximate the unknown value function using the output of the RBF neural network as

$$V(x) = \tfrac{1}{2}(x - x_0)^T \bar{P} (x - x_0) + \tfrac{1}{2}(y - y_0)^T (y - y_0) \qquad (9)$$

where $y_0 = w^T z(x_0, c, \sigma)$, x_0 is the equilibrium point and $\bar{P}$ is the positive definite solution of the Riccati equation corresponding to the linearized approximation of the nonlinear system (1). It should be noted that

$$y - y_0 = w^T \varphi(x, c, \sigma) \qquad (10)$$

where $\varphi(x, c, \sigma) = z(x, c, \sigma) - z(x_0, c, \sigma)$. More specifically, noting that x_0=0 and using (10), the value function (9) can be expressed as

$$V(x) = \tfrac{1}{2} x^T \bar{P} x + \tfrac{1}{2} \varphi^T(x, c, \sigma) w w^T \varphi(x, c, \sigma) \qquad (11)$$

Selecting $V(x)$ as above, it can be verified that $V(x) > 0$ for all $x \in \Re^n, x \neq 0$, and that $V(0) = 0$. Clearly, if the RBFNN parameter vectors w, c, and σ can be found so that the proposed approximation of the value function in (11) satisfies the GHJB equation (5), then $V(x)$ will satisfy the cost function (2) and its time derivative will be given by

$$\dot{V} = -x^T Qx - u^T Ru \qquad (12)$$

which is negative definite. But, since $V(x)$ is positive definite and radially unbounded, this indicates the stability of the closed-loop system. A proper parameter estimation technique, however, is required to find the unknown parameters of the RBFNN for correct approximation of the value function $V(x)$.

The gradient of the value function (11) can be written as

$$\nabla V = \frac{\partial V(x)}{\partial x} = x^T \bar{P} + \varphi^T(x, c, \sigma)(w w^T) \nabla \varphi(x, c, \sigma) \qquad (13)$$

where $\nabla \varphi(x, c, \sigma) = \frac{\partial \varphi(x, c, \sigma)}{\partial x}$. Substituting for ∇V, from the above, in the GHJB equation (5), we get

$$v(\nabla V, u) = \left(x^T \bar{P} + \varphi^T(x, c, \sigma)(w w^T) \nabla \varphi(x, c, \sigma)\right) \left(f(x) + g(x)u(x)\right) + x^T Q x + u^T(x) R u(x) \qquad (14)$$

The above equation can be equivalently written as

$$\xi = h(t,\theta) + v \tag{15}$$

where $\xi = -x^T Qx - u(x)^T Ru(x) - x^T \overline{P}(f(x)+g(x)u(x))$ is a known measurable function of x and u, $h(t,\theta) = \varphi^T(x,c,\sigma)(ww^T)\nabla\varphi(x,c,\sigma)\left(f(x)+g(x)u(x)\right)$ is a known function of the unknown RBFNN parameter vector $\theta = [w^T, c^T, \sigma]^T$, and $v = v(\nabla V, u)$ is the equation error due to approximation of the value function. Clearly, equation (15) is nonlinear in terms of the unknown parameter vector θ. Hence, in order to estimate the unknown RBFNN parameter vector θ, one must use nonlinear parameter estimation techniques. This will be explained in more detail in the following section.

IV. Neural network training

Estimating the RBFNN parameters from equation (15) can be viewed as state estimation problem for an associated parameter system where the RBFNN parameters (weights, centers and widths) are the unknown states to be estimated. It is known that Kalman filters (KF) can be successfully used in the state estimation problem for linear systems. However, because of the nonlinearities in the RBFNN formulation, instead of the KF method an extended Kalman filter (EKF) must be used for state (parameter) estimation. Realizing that the unknown state vector (parameter vector θ) is constant, and using equation (15), the associated parameter dynamics is given as [15]

$$\begin{aligned} \dot{\theta} &= \omega \\ \xi &= h(t,\theta) + v \end{aligned} \tag{16}$$

where θ and ξ are, respectively, the unknown states and measurable outputs of this system, and ω and v are white noise disturbances, with covariance matrices Q_f and R_f, affecting the states and outputs. The EKF algorithm is to estimate the unknown states θ of the system and is given as

$$\dot{\hat{\theta}} = K_f\left(\xi - h(t,\hat{\theta})\right) \tag{17}$$

where the filter gain matrix K_f is given as

$$K_f = SH^T R_f^{-1} \tag{18}$$

where $H^T = \frac{\partial h(\theta)}{\partial \theta}\Big|_{\theta=\hat{\theta}}$ is the observation matrix for the linearized model of the system (16) and $S > 0$ is a symmetric positive definite matrix, which satisfies the following filter differential Riccati equation [13], i.e.,

$$\dot{S} = 2\alpha S + Q_f - SHR_f^{-1}H^T S \tag{19}$$

where $S(t_0) = S_0 = S_0^{\,T} > 0$, $Q_f \geq 0$, $R_f > 0$, and $\alpha > 0$. The convergence properties of the EKF are explained in [16]. However, the region of convergence for the EKF algorithm in this case may be small. Here, to improve the region of convergence, an adaptive extended Kalman filter (AEKF) is proposed as

$$\dot{\hat{\theta}} = K_f\,\tilde{\xi} \tag{20}$$

where $\tilde{\xi} = \xi - \hat{\xi}$, $\hat{\xi} = h(t,\hat{\theta})$, and the filter gain matrix K_f is given by

$$K_f = S\hat{C}^T R_f^{-1} \tag{21}$$

Also, the corresponding adaptive output matrix $\hat{C}$ is adjusted as

$$\dot{\hat{C}} = \left[\lambda\tilde{\xi} + \gamma sign\left(\tilde{\xi}\right)\right]\hat{\theta}^T \tag{22}$$

$$\gamma = \gamma_0 + \gamma_1\int_0^t \tilde{\xi}^2 \tag{23}$$

where $\lambda > 0$, $\hat{C}(0) = \hat{H}(0) = \dfrac{\partial\hat{h}(0)}{\partial\hat{\theta}}$, $\gamma_0, \gamma_1 > 0$, $\alpha > 0$, $sign(\tilde{\xi}) \cong sat\left(\frac{\tilde{\xi}}{\varepsilon}\right)$ and $\varepsilon << 1$. Moreover, the symmetric positive definite matrix, $S > 0$, satisfies a filter differential Riccati equation (FDRE) [16], given as

$$\dot{S} = 2\alpha S - S\hat{C}^T R_f^{-1}\hat{C}S + Q_f \tag{24}$$

where $S(0) = S^T(0) = S_0 > 0$. The convergence of the RBFNN parameters (weights, centers, and widths) is shown in the following lemma.

Lemma:

Consider the dynamics of the RBFNN parameters (16). If these unknown parameters are estimated according to the proposed AEKF algorithm (20)-(24), then the output and the parameter errors, $\tilde{\xi}$, $\tilde{\theta}$, converge to zero, asymptotically.

Proof:

Let $\tilde{\theta} = \theta - \hat{\theta}$. Then

$$\dot{\tilde{\theta}} = -K_f\,\tilde{\xi} + \omega \tag{25}$$

Using power series expansion, let us express $h = h(t,\theta)$ as

$$h = h_0 + C(t,\theta)\,\theta \tag{26}$$

where $h_0 = h(t,0)$, and $C(t,\theta)$ is a nonlinear vector. The estimate of $h(t,\theta)$ can then be written as

$$\hat{h} = h_0 + \hat{C}(t,\hat{\theta})\,\hat{\theta} \tag{27}$$

where $\hat{h} = h(t,\hat{\theta})$. Now define $\tilde{h} = h - \hat{h} = C\,\theta - \hat{C}\,\hat{\theta}$. Then

$$\begin{aligned} \dot{\tilde{h}} &= \dot{h} - \dot{\hat{h}} = C\,\dot{\theta} + \dot{C}\,\theta - \hat{C}\,\dot{\hat{\theta}} - \dot{\hat{C}}\,\hat{\theta} \\ &= C\,\omega + \dot{C}\,\theta - \hat{C}\,S\,\hat{C}^T R_f^{-1}\tilde{\xi} - \dot{\hat{C}}\,\hat{\theta} \end{aligned} \tag{28}$$

Then, noting that $\tilde{\xi} = \tilde{h} + v$ and using $\dot{\hat{C}}$ form (22)-(23), one can write

$$\dot{\tilde{\xi}} = \mu_1 - \hat{C}\,S\,\hat{C}^T R_f^{-1}\tilde{\xi} - \left(\lambda\tilde{\xi} + \gamma sign(\tilde{\xi})\right)\left\|\hat{\theta}\right\|^2 \tag{29}$$

where $\mu_1 = C\,\omega + \dot{C}\,\theta + \dot{v}$ is bounded. Now, using the Lyapunov function $V_{\tilde{\xi}} = \tilde{\xi}^2$ and its derivative $\dot{V}_{\tilde{\xi}} \leq -\eta_0\,\tilde{\xi}^2$, with $\eta_0 = 2\left(\lambda\left\|\hat{\theta}\right\|_{\min}^2 + \frac{\eta}{|\tilde{\xi}|_{\max}}\right)$, $\eta > 0$ and $\gamma\left\|\hat{\theta}\right\|_{\min}^2 - |\mu_1|_{\max} \geq \eta$, one can show that $\tilde{\xi}$ converges to zero. However, since $C\theta + v$ is bounded and $\tilde{\xi} = \tilde{h} + v = C\theta - \hat{C}\hat{\theta} + v$ goes to zero, $\hat{C}\hat{\theta}$ must be bounded and must converge to $C\theta + v$.

Moreover, since $V_{\hat{C}} = \left\|\hat{C}\right\|^2 = \hat{C}\,\hat{C}^T$ is lower bounded and its derivative $\dot{V}_{\hat{C}} = 2\hat{C}\dot{\hat{C}}^T = 2\hat{C}\,\hat{\theta}\left[\lambda\tilde{\xi} + \gamma sign\left(\tilde{\xi}\right)\right]$ converges to zero, then $V_{\hat{C}}$ and hence $\left\|\hat{C}\right\|$ must be bounded. But, since $\hat{C}\hat{\theta}$ is bounded, then $\left\|\hat{\theta}\right\|$ must also be bounded, Now consider a Lyapunov function candidate

$$V = \tilde{\theta}^T \Pi\,\tilde{\theta} + \tilde{\xi}^2 \tag{30}$$

where $\Pi = \Pi^T > 0$ such that $S = \Pi^{-1}$ and $\dot{\Pi} = -\Pi\,\dot{S}\,\Pi$. Take the time-derivative of $V(x)$, apply the Young's inequality to the terms $2\tilde{\theta}^T\Pi\omega$ and $-2\,\tilde{\theta}^T\hat{C}^T R_f^{-1}\tilde{\xi}$, add and subtract the term $2\alpha\tilde{\theta}^T\Pi\,\tilde{\theta} + 2\tilde{\theta}^T\hat{C}^T R_f^{-1}\hat{C}\tilde{\theta}$, note that $\tilde{\theta}^T\hat{C}^T R_f^{-1}\hat{C}\tilde{\theta} \le \frac{\|\hat{C}\|_{\max}^2 \|S\|_{\max}}{R_f}\tilde{\theta}^T\Pi\,\tilde{\theta}$, and substitute for $\dot{\tilde{\xi}}$ from (29). Then, we get

$$\begin{aligned}
\dot{V} &= \tilde{\theta}^T\dot{\Pi}\,\tilde{\theta} + 2\,\tilde{\theta}^T\Pi\,\dot{\tilde{\theta}} + 2\,\tilde{\xi}\dot{\tilde{\xi}} \\
&= -\tilde{\theta}^T\Pi\,\dot{S}\,\Pi\,\tilde{\theta} + 2\tilde{\theta}^T\Pi\omega - 2\,\tilde{\theta}^T\hat{C}^T R_f^{-1}\tilde{\xi} + 2\,\tilde{\xi}\dot{\tilde{\xi}} \\
&\le -\tilde{\theta}^T\Pi\,\dot{S}\,\Pi\,\tilde{\theta} + 2\alpha_1\tilde{\theta}^T\Pi\,\tilde{\theta} + \tfrac{\|\Pi\|_{\max}}{2\alpha_1}\|\omega\|^2 \\
&\quad + 2\,\alpha_2\,\tilde{\theta}^T\Pi\,\tilde{\theta} + 2\,\tilde{\xi}\left[\dot{\tilde{\xi}} + \tfrac{\hat{C}S\hat{C}^T}{4\alpha_2 R_f^2}\tilde{\xi}\right] \\
&\le \tilde{\theta}^T\Pi\left[-\dot{S} + 2\alpha S - 2S\hat{C}^T R_f^{-1}\hat{C}S\right]\Pi\,\tilde{\theta} + \tfrac{\|\Pi\|_{\max}}{2\alpha_1}\|\omega\|^2 \\
&\quad + 2\left(\alpha_1 + \alpha_2 + \tfrac{\|\hat{C}\|_{\max}^2\|S\|_{\max}}{R_f} - \alpha\right)\tilde{\theta}^T\Pi\,\tilde{\theta} \\
&\quad + 2\,\tilde{\xi}\left[\mu_1 - \tfrac{(4\alpha_2 R_f - 1)\hat{C}S\hat{C}^T}{4\alpha_2 R_f^2}\tilde{\xi} - \left(\lambda\tilde{\xi} + \gamma sign(\tilde{\xi})\right)\left\|\hat{\theta}\right\|^2\right]
\end{aligned} \tag{31}$$

Now substitute for $\dot{S}$ from (24) in the above, and note that since $Q_f \ge 0$ and $R_f > 0$, the solution S to the FDRE (24) is bounded both from above and below. Let $\alpha_1 > 0$, $\alpha_2 \ge \frac{1}{4R_f}$, $\alpha \ge \alpha_1 + \alpha_2 + \frac{\|\hat{C}\|_{\max}^2\|S\|_{\max}}{R_f}$. Then, we get

$$\begin{aligned}
\dot{V} &\le -\tilde{\theta}^T\Pi\left[Q_f + S\,\hat{C}^T R_f^{-1}\hat{C}\,S\right]\Pi\,\tilde{\theta} + \tfrac{\|\Pi\|_{\max}}{2\alpha_1}\|\omega\|^2 \\
&\quad + 2\,\tilde{\xi}\left[\mu_1 - \left(\lambda\tilde{\xi} + \gamma sign(\tilde{\xi})\right)\left\|\hat{\theta}\right\|^2\right]
\end{aligned} \tag{32}$$

However, as long as $\tilde{\xi} \ne 0$, γ will increase and, since $\left\|\hat{\theta}\right\| \ne 0$, after a finite time we have $\gamma\left\|\hat{\theta}\right\|_{\min}^2 - |\mu_1|_{\max} \ge \eta$ for some $\eta > 0$. Now, let $\eta_0 = 2\left(\lambda\left\|\hat{\theta}\right\|_{\min}^2 + \frac{\eta}{|\tilde{\xi}|_{\max}}\right)$ with $\eta > 0$. Also let $\lambda_0 = \min\left\{\left\|\Pi Q_f \Pi\right\|_{\min}, \eta_0\right\}$ and choose $\alpha_1 > 0$ such that $\alpha_1 \ge \frac{\|\Pi\|_{\max}}{2\delta_0}\|\omega\|_{\max}^2$ for some $\delta_0 > 0$. Then, the derivative of the Lyapunov function will become

$$\dot{V} \le -\lambda_0\left(\left\|\tilde{\theta}\right\|^2 + \left|\tilde{\xi}\right|^2\right) + \delta_0 \tag{33}$$

This proves the boundedness of the Lyapunov function V as well as $\left\|\tilde{\theta}\right\|$ and $\left|\tilde{\xi}\right|$. Moreover, let us define the residual set $r_0 = \left\{(\tilde{\theta}, \tilde{\xi}) \middle| \left(\left\|\tilde{\theta}\right\|^2 + \left|\tilde{\xi}\right|^2\right) \le \varepsilon_0\right\}$ where $\varepsilon_0 = \frac{\delta_0}{\lambda_0} > 0$. Then, it is easy to show that the total estimation error $\left\|\tilde{\theta}\right\|^2 + \left|\tilde{\xi}\right|^2$ must converge to the small residual set r_0. Furthermore, in case $\omega = 0$, we will have $\varepsilon_0 = 0$ and $r_0 = \{0\}$, and hence V and both $\left\|\tilde{\theta}\right\|$ and $\left|\tilde{\xi}\right|$ must converge to zero, asymptotically.

Given $Q_f \ge 0$ and $R_f > 0$, and assuming that the associated parameter system is observable, the FDRE (24) can be solved for the covariance matrix $S > 0$ continuously and the time varying gain K_f can be determined from equation (21). The RBFNN parameters are then estimated using equation (20) with the knowledge of K_f and measurement of system states x. Assuming the nonlinear system is controllable, with the estimated RBFNN parameters, the nonlinear system can be controlled using the optimal control (3) via the solution of the GHJB equation (5). It can be shown that the closed-loop system, consisting of the AEKF parameter estimator and the GHJB-based optimal control working together, simultaneously, is globally stable and that both the estimation and the control errors will converge to a small residual set.

V. Simulation Results

In this section, SISO nonlinear systems are considered for regulation problem. The proposed technique is applied for the optimal control design. Two examples are presented.

Example 1:

Consider the following first order nonlinear system

$$\dot{x} = x^3 + u \tag{34}$$

The objective is to design the optimal control for this system so as to minimize the cost $J = \int_0^\infty (x^2 + 2x^4 + u^2)dt$.

In this example we compare our proposed neural network based optimal control with the exact optimal control. The exact optimal control can be found analytically by solving the HJB equation [3], which results in

$$u^*(x) = -\frac{f(x)}{b} - sign(bx)\sqrt{\frac{f^2(x)}{b^2} + \frac{l(x)}{r}} \tag{35}$$

For the above example, the optimal control is given by

$$u^*(x) = -x - 2x^3 \tag{36}$$

For our proposed method the number of basis functions and, hence, the number of centers, widths and weights were selected to be 5 and the neural network is trained by the adaptive EKF algorithm. Fig.2 shows the system states. Fig.3 shows a comparison of the value functions corresponding to the exact optimal control found in equation (36) and that of our proposed method. The comparison establishes that our proposed method is nearly optimal. Also, Fig.4 shows the norm of the RBFNN parameter vectors $\hat{w}$,

$\hat{c}$ and $\hat{\sigma}$. It can be seen that these parameters converge to some constant values, quickly.

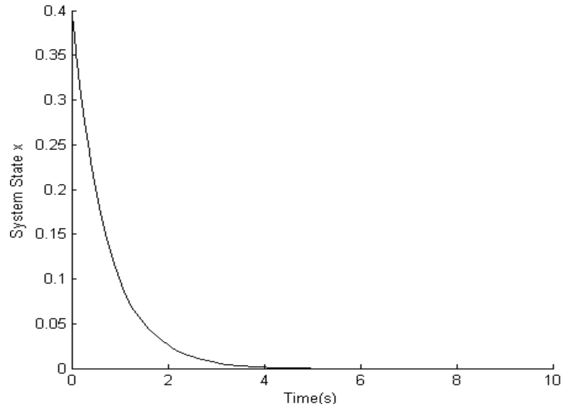

Fig.2- System state

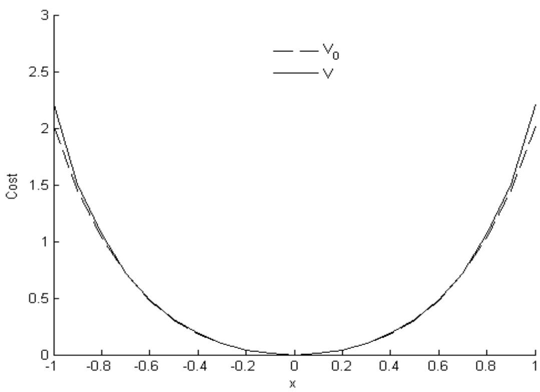

Fig. 3- Performance cost

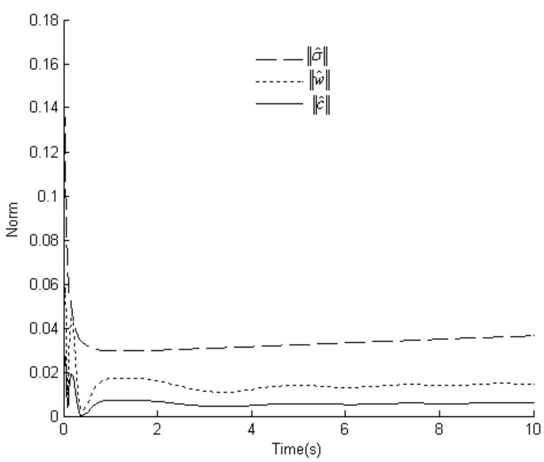

Fig. 4- Norms of the RBFNN parameter vectors

Example 2:

Consider the following 2nd order SISO nonlinear system

$$\dot{x} = f(x) + g(x)u = \begin{bmatrix} -x_1^3 - x_2 \\ x_1 + x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (37)$$

where the state vector x is assumed to be completely known. The goal is to find the optimal control so as to minimize the performance cost

$$J = \int_0^\infty [x^T Q x + u^T R u] dt \quad (38)$$

where Q and R are the positive definite matrices and chosen as R=1 and $Q = I_{2\times 2}$ for this problem. It can be verified that the system has an equilibrium point at $x_0 = (0,0)^T$. Again, the value function $V(x)$ is approximated, using the output of an RBF neural network, as

$$V(x) = \tfrac{1}{2} x^T \overline{P} x + \tfrac{1}{2} \varphi^T(x) w\, w^T \varphi(x) \quad (39)$$

The positive definite matrix $\overline{P}$ of the above value function is obtained by linearizing the system (37), around its equilibrium point x=0, and by solving the following Riccati equation, i.e.,

$$A^T \overline{P} + \overline{P} A + Q - \overline{P} B R^{-1} B^T \overline{P} = 0 \quad (40)$$

where $A = \frac{\partial F}{\partial x}\Big|_{\substack{x=0\\u=0}}$, $B = \frac{\partial F}{\partial u}\Big|_{\substack{x=0\\u=0}}$.

The number of radial basis functions selected for this problem is 8. The weights, centers, and widths of the RBFNN are estimated, form the GHJB equation, using the proposed AEKF algorithm explained in the previous section. The optimal control law was then found from the estimated value function, based on the proposed method. In addition, for comparative study, two other approximation techniques, namely the exact feedback linearization (FL) and successive Galerkin approximation (SGA) were used to find the optimal control. Consequently, the proposed optimal control, the FL control and the SGA control were each applied to the system, independently, under the same conditions. The exact feedback linearization control law u_{FL} for this SISO nonlinear system was adopted from [14], which is given as

$$u_{FL} = 3x_1^5 + 3x_1^2 x_2 - x_2 + 0.4142x_1 - 1.3522(x_1^3 + x_2) \quad (41)$$

The control law obtained from the SGA method, using 8 basis functions, is given as [14]

$$\begin{aligned} u_{SGA} &= -0.4215x_1 - 2.2225x_2 - 0.4784x_1^3 \\ &\quad + 0.2719x_1^2 x_2 + 0.6494x_1 x_2^2 + 0.0588x_2^3 \end{aligned} \quad (42)$$

All the simulations were carried out using MATLAB/ SIMULINK. Fig.5 shows the system states x_1 and x_2 for the proposed control. A comparison of the value function found using these control methods with the initial conditions $(x_1\ 0)^T, x_1 \in [-1,1]$ is shown in Fig.6. In this figure V_{FL} is the value function associated with the control law u_{FL}, V_{SGA} is the value function associated with the control law u_{SGA} and V is the value function associated with the proposed control law based on the RBF neural network. From these figures it is clear that the proposed control method provides a better approximation of the optimal control as compared to those found in [17] for the same number of basis functions. It is shown in [17] that the controller with 15 basis functions performs better than the controller with 8 basis functions. With 15 basis functions, the proposed method performs slightly better than 8 basis functions and is similar to [17] with 15 basis functions. Fig.7

shows the norm of the RBFNN parameter vectors $\hat{w}$, $\hat{c}$ and $\hat{\sigma}$, where they all converge to some constant values.

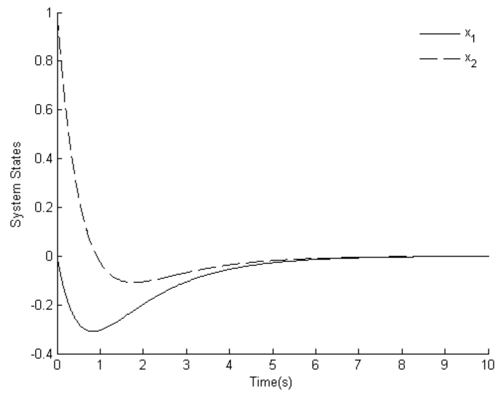

Fig. 5- System states

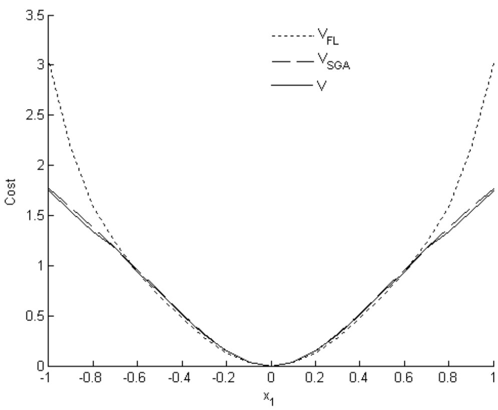

Fig. 6- Performance cost

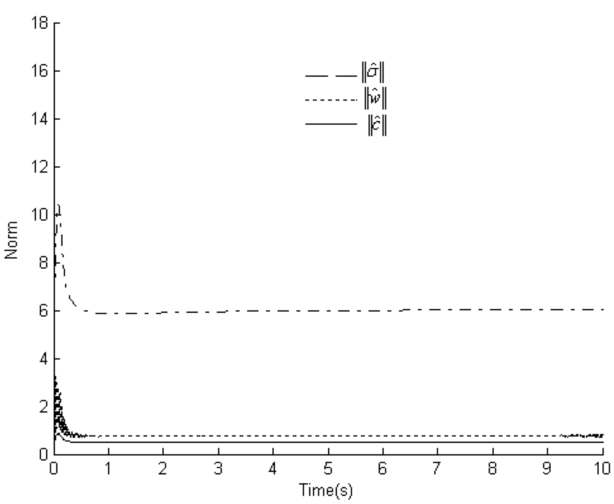

Fig. 7- Norms of the RBFNN parameter vectors

VI. CONCLUSION

This paper presents an approach to finding the optimal control law for nonlinear systems using neural networks. The design procedure approximately solves the generalized Hamilton-Jacobi-Bellman (GHJB) equation by estimating the value function using nonlinear radial basis function neural networks (RBFNN). The neural network is trained by an adaptive extended Kalman filter (AEKF), which estimates the RBFNN parameters, online. The proposed nonlinear optimal control method was applied to a 1st and a 2nd order SISO nonlinear system. In the case of 1st order nonlinear system the proposed approximation of the optimal control was compared with the exact optimal control to show that the proposed method indeed generates a nearly optimal control.

Acknowledgement

Partial support by the NSF grant No. 0115507 is gratefully acknowledged.

REFERENCES

[1] Peter Dorato. Linear quadratic control, an introduction. Krieger publishing company: Florida, 2000

[2] Curtis P. Mracek and James R. Cloutier. Control designs for the nonlinear benchmark problem via the state dependent Riccati equation method. *International journal of robust and nonlinear control*, 1998; 8(4-5): 401-433

[3] W. L. Garrard and J.M.Jordan. Design of nonlinear automatic flight control systems. *Automatica*, 1977; 13: 497-505

[4] W.L.Garrard, D.F.Ennis and S.A. Snell. Nonlinear feedback control of highly maneuverable aircraft *International journal of control*, 1992; 56(4): 799-812

[5] G.Saridis, C.S.Lee. An approximation theory of optimal control for trainable manipulators. *IEEE Transactions on Systems, Man, Cybernetics*, 1979; 9(3):152-159.

[6] R.Beard. Improving the closed-loop performance of nonlinear systems. PhD Thesis, Rensselaer Polytechnic Institute, Troy, New York, 1995.

[7] C Park, P Tsiotras. Sub-optimal feedback control using a successive-Galerkin algorithm. *American control conference* 2003; 3: 1926-1931

[8] S. Kumar, R. Padhi and L. Behra " Continuous-time Single Network Adaptive Critic for Regulator Design of Nonlinear Control Affine Systems" *Proceedings of 17th World Congress of the International Federation of Automatic Control*, 2008, Seol, S. Korea.

[9] F. Pourboghrat, H.Pongpairoj, Ziqian Liu, F.Farid. Dynamic neural Networks for modeling and control of nonlinear systems. *Intelligent Automation and soft computing,* 2003; 9(2): 61-70.

[10] Chen, F.C., Liu C. Adaptively controlling nonlinear continuous-time systems using multilayer neural networks. *IEEE Transactions on Automatic Control,* 1994; 39(6): 1306-1310.

[11] FL Lewis, M Abu-Khalaf. A Hamilton-Jacobi setup for constrained neural network control. *Proceedings of the IEEE International Symposium on Intelligent Control* 2003: 1-15.

[12] M.S.Ahmed "Neural controllers for nonlinear state feedback L_2- gain control" *IEE Proceedings of Control Theory Applications*, 2000 147(3).

[13] J. K. Sing., D. K. Basu, and M. Nasipuri " Improved K- means algorithm in the design of radial basis function network. *Proceedings of Convergent Technologies for Asia-Pacific Region*, Tencon.2, 841-845.

[14] J. J. Murray, C. J. Cox, G. G. Lendaris and R. Saeks " Adaptive Dynamic programming" *IEEE Trans. On Systems, Man and Cybernetics, Part C.*, Vol. 32, No. 2, May 2002.

[15] Dan Simon. Training radial basis neural networks with the extended kalman filter. *Neurocomputing* 2001.

[16] K. Reif, F. Sonnemann and R. Unbehauen. An EKF-based nonlinear observer with a prescribed degree of stability. *Automatica* 1998 34(9) :1119-1123.

[17] R. Beard, G. Saridis and J. Wen. Galerkin approximation of the generalized Hamilton-Jacobi-Bellman equation. *Automatica* 1997 33(12): 2159-2177.