REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE:	3. REPORT TYPE AND DATES COVERED New Reprint 20-Jun-2005 - 1-Jun-2006

4. TITLE AND SUBTITLE A Neural Network Solution for Fixed-Final Time Optimal Control of Nonlinear Systems	5. FUNDING NUMBERS W911NF0510314
6. AUTHORS Tao Cheng, Frank L. Lewis, and Murad Abu-Khalaf	8. PERFORMING ORGANIZATION REPORT NUMBER
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of Texas - Arlington Grants and Contracts Services Box 19145 Arlington, TX 76019 -0145	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSORING / MONITORING AGENCY REPORT NUMBER 46972-CI.4

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12. DISTRIBUTION AVAILIBILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for Public Release; Distribution Unlimited	

13. ABSTRACT (Maximum 200 words)

The abstract is below since many authors do not follow the 200 word limit

14. SUBJECT TERMS	15. NUMBER OF PAGES
fixed final time control, optimal control, neurocontrol, finite time control	Unknown due to possible attachments
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION ON THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev .2-89)
Prescribed by ANSI Std.
239-18 298-102

Report Title

A Neural Network Solution for Fixed-Final Time Optimal Control of Nonlinear Systems

ABSTRACT

We consider the use of neural networks and Hamilton-Jacobi-Bellman equations towards obtaining fixed-final time optimal control laws in the input nonlinear systems. The method is based on Kronecker matrix methods along with neural network approximation over a compact set to solve a time-varying Hamilton-Jacobi-Bellman equation. The result is a neural network feedback controller that has time-varying coefficients found by a priori offline tuning. Convergence results are shown. The results of this paper are demonstrated on two examples.

Proc. Mediterranean Conf. Control and Automation
paper TM2-3, Ancona, June 2006.

A Neural Network Solution for Fixed-Final Time Optimal Control of Nonlinear Systems

Tao Cheng, Frank L. Lewis, *Fellow, IEEE*, and Murad Abu-Khalaf

***Abstract*—We consider the use of neural networks and Hamilton-Jacobi-Bellman equations towards obtaining fixed-final time optimal control laws in the input nonlinear systems. The method is based on Kronecker matrix methods along with neural network approximation over a compact set to solve a time-varying Hamilton-Jacobi-Bellman equation. The result is a neural network feedback controller that has time-varying coefficients found by a priori offline tuning. Convergence results are shown. The results of this paper are demonstrated on two examples.**
Keywords: Finite-horizon optimal control, Hamilton-Jacobi-Bellman, Neural Network control

I. INTRODUCTION

The solution of the Hamilton-Jacobi-Bellman (HJB) equation resulting in finite-time optimal control laws for nonlinear systems is a challenging problem. It is known that this optimization problem [16], requires solving a time-varying HJB equation that is hard to solve in most cases. Approximate HJB solutions have been confronted using many techniques such as those developed by Saridis and Lee [27], Beard et. Al [4][5][6], Beard, Bertsekas and Tsitsiklis [7], Munos et. al [22], Kim, Lewis and Dawson [14], Liu and Balakrishnan [17], Lyshevski and Meyer [20] and Lyshevski [18][19]. Huang and Lin [13] provided a Taylor series expansion of the HJI equation which is closely related to the HJB equation.

Successful neural networks (NN) controllers not based on optimal techniques have been reported in Chen and Liu [8], Lewis, Jagannathan and Yesildirek [15], Polycarpou [24], Rovithakis and Christodoulou [25], Sanner and Slotine [26], Ge [11]. It has been shown that NN can effectively extend adaptive control techniques to nonlinearly parameterized systems. NN applications to an optimal control via the HJB equation were first proposed by Werbos [21]. Parisini and Zoppoli [23] used NN to derive optimal control laws for discrete-time stochastic nonlinear systems.

In this paper, we use NN to approximately solve the time-varying HJB equation employing a nonquadratic functional. It is shown that using a NN approach, one can simply transform the problem into solving an ordinary differential equation (ODE) equation backwards in time. The coefficients of this ODE are obtained by the weighted residuals method.

Motivated by the important results in [4], we are able to approximately solve the time-varying HJB equation without policy iteration using the so-called GHJB equation followed by control law updates. We accomplish this by using a neural network approximation for the value function which is based on a universal basis set.

Manuscript received January 23, 2006. This work was supported by NSF ECS-0501451 and ARO W91NF-05-1-0314.
Tao Cheng, Frank L. Lewis and M. Abu-Khalaf are with the Automation and Robotics Research Institute, The University of Texas at Arlington, TX 76118, USA. (Tel./fax: +1 817 272 5938, E-mail: chengtao@arri.uta.edu)

II. PROBLEM STATEMENT

Consider an affine in the control nonlinear dynamical system of the form

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

where $x \in \Re^n$, $f(x) \in \Re^n$, $g(x) \in \Re^{n\times m}$ and the input $u(t) \in R^m$. The dynamics $f(x)$ and $g(x)$ are assumed to be known. Assume that $f + gu$ is Lipschitz continuous on a set $\Omega \subseteq \Re^n$ containing the origin, and that system (1) is stabilizable in the sense that there exists a continuous control on Ω that asymptotically stabilizes the system. It is desired to find the control u that minimizes a generalized nonquadratic functional

$$V(x(t_0),t_0) = \phi(x(t_f),t_f) + \int_{t_0}^{t_f} [Q(x) + W(u)]dt \tag{2}$$

with $Q(x)$, $W(u)$ positive definite on Ω, i.e. $\forall x \neq 0$, $x \in \Omega$, $Q(x) > 0$ and $x = 0 \Rightarrow Q(x) = 0$. A common choice for $W(u) = u^T Ru$, where $R > 0$. The final time t_f is fixed.

Definition 1. Admissible Controls.

A control u is defined to be admissible with respect to (2) on Ω, denoted by $u \in \Psi(\Omega)$, if u is continuous on Ω, $u(0) = 0$, u stabilizes (1) on Ω, and $\forall x_0 \in \Omega$, $V(x_0)$ is finite. ■

Definition 2. Sobolev Space.

$H^{m,p}(\Omega)$: Let Ω be an open set in $\Re^n$ and let $u \in C^m(\Omega)$. Define a norm on u by

$$\|u\|_{m,p} = \sum_{0\le|\alpha|\le m} \left(\int_\Omega \left|D^\alpha u(x)\right|^p dx\right)^{1/p}, \quad 1 \le p < \infty .$$

This is the Sobolev norm in which the integration is Lebesgue. The completion of $u \in C^m(\Omega)$: $\|u\|_{m,p} < \infty$ with respect to $\|\cdot\|_{m,p}$ is the Sobolev space $H^{m,p}(\Omega)$. For $p = 2$, the Sobolev space is a Hilbert space. ■

The convergence proofs of the least-squares method are done in the Sobolev function space $H^{1,2}(\Omega)$ setting [2], since we require to prove the convergence of both $V_L(x)$ and its gradient.

An infinitesimal equivalent to (2) is [16]

$$-\frac{\partial V}{\partial t}=L+\left(\frac{\partial V}{\partial x}\right)^T (f+gu). \tag{3}$$

This is a time-varying partial differential equation. It is in fact a Lyapunov equation that yields the value V for any given u and is solved backward in time from $t=t_f$. By setting $t_0=t_f$ in (2) its boundary condition is seen to be

$$V(x(t_f),t_f)=\phi(x(t_f),t_f). \tag{4}$$

According to Bellman's optimality principle [16], the optimal cost is given by

$$-\frac{\partial V^*}{\partial t}=\min_{u(t)}\left(L+\left(\frac{\partial V^*}{\partial x}\right)^T f\right), \tag{5}$$

which yields the optimal control.

$$u^*(x)=-\frac{1}{2}R^{-1}g^T\frac{dV^*}{dx}, \tag{6}$$

where $V^*(x)$ is the optimal value function. Substituting (6) into (5) yields the well-known time-varying HJB equation [16]

$$\begin{aligned}HJB(V^*)&=\frac{\partial V^*}{\partial t}+\frac{\partial V^*}{\partial x}f+Q\\&-\frac{1}{4}\frac{\partial V^{*T}}{\partial x}gR^{-1}g^T\frac{\partial V^*}{\partial x}=0\end{aligned}. \tag{7}$$

This equation and (6) provide the solution to fixed-final time optimal control for general nonlinear systems. However, this equation is generally impossible to solve.

III. Nonlinear Fixed-Final-Time HJB Solution by NN Least-Squares Approximation

The HJB equation (7) is difficult to solve for the cost function $V(x)$. In this section, neural networks are used to solve approximately for the value function in (7) over Ω by approximating the cost function $V(x)$. The result is an efficient, practical, and computationally tractable solution algorithm to find nearly optimal state feedback controllers for nonlinear systems.

A *NN Approximation of* $V(x)$

It is well known that NN can be used to approximate smooth functions on prescribed compact sets (Hornik [12]). Since the analysis required here is restricted to the region of asymptotically stable (RAS) of some initial stabilizing controller, NN are natural for this application. We use the following equation to approximate V

$$V_L(x)=\sum_{j=1}^{L}w_j\sigma_j(x)=w_L^T(t)\sigma_L(x), \tag{8}$$

which is a NN with activation functions $\sigma_j(x)\in C^1(\Omega)$, $\sigma_j(0)=0$. The NN weights are $w_j(t)$ and L is the number of hidden-layer neurons. $\boldsymbol{\sigma}_L(x)\equiv[\sigma_1(x)\sigma_2(x)...\sigma_L(x)]^T$ is the vector of activation function, $\mathbf{w}_L(t)\equiv[w_1(t)w_2(t)...w_L(t)]^T$ is the vector of NN weights.

Since one requires $\partial V/\partial t$ in (7), the NN weights are selected to be time-varying. This is similar to methods such as assumed mode shapes in the study of flexible mechanical systems [3]. However, here $\boldsymbol{\sigma}_L(x)$ is a NN activation vector, not a set of eigenfunctions. That is, the NN approximation property significantly simplifies the specification of $\boldsymbol{\sigma}_L(x)$. For the infinite final time case, the NN weights are constant [1]. The NN weights will be selected to minimize a residual error in a least-squares sense over a set of points sampled from a compact set Ω inside the RAS of the initial stabilizing control [10].

Note that

$$\frac{\partial V_L}{\partial x}=\frac{\partial\boldsymbol{\sigma}_L^T}{\partial x}\mathbf{w}_L(t)\equiv\nabla\boldsymbol{\sigma}_L{}^T\mathbf{w}_L(t), \tag{9}$$

where $\nabla\boldsymbol{\sigma}_L$ is the Jacobian $\partial\boldsymbol{\sigma}_L/\partial x$, and that

$$\frac{\partial V_L}{\partial t}=\dot{\mathbf{w}}_L^T(t)\boldsymbol{\sigma}_L(x). \tag{10}$$

Therefore approximating $V(x)$ by $V_L(x)$ in the HJB equation (7) results in

$$\begin{aligned}&-\dot{\mathbf{w}}_L^T(t)\boldsymbol{\sigma}_L(x)-\mathbf{w}_L^T(t)\nabla\boldsymbol{\sigma}_L(x)f(x)\\&+\frac{1}{4}\mathbf{w}_L^T(t)\boldsymbol{\sigma}_L(x)g(x)R^{-1}g^T(x)\boldsymbol{\sigma}_L^T(x)\mathbf{w}_L(t),\\&-Q(x)=e_L(x)\end{aligned} \tag{11}$$

or

$$HJB\left(V_L(x)=\sum_{j=1}^{L}w_j\sigma_j(x),u\right)=e_L(x), \tag{12}$$

where $e_L(x)$ is a residual equation error. From (6) the corresponding control input is

$$u_L(t)=-\frac{1}{2}R^{-1}g^T(x)\nabla\boldsymbol{\sigma}_L{}^T(x)\mathbf{w}_L(t)\cdot \tag{13}$$

To find the least-squares solution for $\mathbf{w}_L(t)$, the method of weighted residuals is used [10]. The weight derivatives $\dot{\mathbf{w}}_L(t)$ are determined by projecting the residual error onto $de_L(x)/d\dot{\mathbf{w}}_L(t)$ and setting the result to zero $\forall x\in\boldsymbol{\Omega}$ using the inner product, i.e.

$$\left\langle\frac{de_L(x)}{d\dot{\mathbf{w}}_L(t)},e_L(x)\right\rangle_\Omega=0. \tag{14}$$

From (11) we can get

$$\frac{de_L(x)}{d\dot{\mathbf{w}}_L}=\boldsymbol{\sigma}_L(x). \tag{15}$$

Therefore one obtains

$$\begin{aligned}&\left\langle-\dot{\mathbf{w}}_L^T(t)\boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega+\left\langle-\mathbf{w}_L^T(t)\nabla\boldsymbol{\sigma}_L(x)f(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega\\&+\left\langle\frac{1}{4}\mathbf{w}_L^T(t)\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1}g^T(x)\nabla\boldsymbol{\sigma}_L^T(x)\mathbf{w}_L(t),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega\\&+\left\langle-Q(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega=0\end{aligned} \tag{16}$$

So that

$$
\begin{aligned}
&\dot{\mathbf{w}}_L(t)= \\
&-\langle \boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega^{-1}\cdot\langle \nabla\boldsymbol{\sigma}_L(x)f(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega\cdot\mathbf{w}_L(t) \\
&+\langle \boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega^{-1}\cdot\left\langle \begin{matrix}\frac{1}{4}\mathbf{w}_L^T(t)\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1} \\ \cdot g^T(x)\nabla\boldsymbol{\sigma}_L^T\mathbf{w}_L(t),\boldsymbol{\sigma}_L(x)\end{matrix}\right\rangle_\Omega \\
&-\langle \boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega^{-1}\cdot\langle Q(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega
\end{aligned}
\qquad (17)
$$

with boundary condition $V(t_f,x)=\phi(x(t_f),t_f)$.

Therefore, the NN weights are simply found by integrating this nonlinear ODE backwards in time.

Following two lemmas show that this procedure provides a nearly optimal solution for the time-varying optimal control problem if time-varying L is selected large enough.

Lemma 1. Convergence of Approximate Value Function. If Ω is compact, $Q(x)$ are continuous on Ω and are in the space $span\{\sigma_j\}_1^\infty$, and if the coefficients $|w_j(t)|$ are uniformly bounded for all L, then $\|V_L-V\|_{L_2(\Omega)}\to 0$ as L increases.

Proof. See [9]. ■

Lemma 2. Convergence of Value Function Gradient. Under the hypothesis of Lemma 1,

$\left\| dV_L/dx - dV/dx \right\|_{L_2(\Omega)} \to 0$ as L increases.

Proof. See [9]. ■

At this point we have proven convergence in the mean of the approximate value function and the value function gradient. This demonstrates convergence in the mean in Sobolev space $H^{1,2}(\Omega)$.

Lemma 3. Admissibility of $u_L(x)$. If the conditions of Lemma 1 are satisfied, then $\exists L_0 : L\ge L_0, u_L\in\Psi(\Omega)$.

Proof. Define

$$V(x,W)=\phi(x(t_f),t_f)+\int_{t_0}^{t_f}[Q(x)+W(u)]dt\,.$$

We must show that for L sufficiently large, $V(x,u_L)<\infty$ when $V(x,u)<\infty$. But $\phi(x(t_f),t_f)$ depends continuously on W, i.e., small variations in W result in small variations in ϕ. Also since $\|u_L(\cdot)\|^2_{L_2(\Omega)}$ can be made arbitrarily close to $\|u(\cdot)\|^2_{L_2(\Omega)}$, $V(x,u_L)$ can be made arbitrarily close to $V(x,u)$. Therefore for L sufficiently large, $V(x,u_L)<\infty$ and hence $u_L(x)$ is admissible. ■

Lemma 3 shows that if the number L of hidden layer units is large enough, the proposed solution method yields an admissible control.

B *Optimal Algorithm Based on NN Approximation*

Solving the integration in (16) is expensive computationally. Since evaluation of the L_2 inner product over Ω is required. This can be addressed using the collocation method [10]. The integrals can be well approximated by discretization. A mesh of points over the integration region can be introduced on Ω of size Δx. The terms of (16) can be rewritten as follows

$$A=\left[\boldsymbol{\sigma}_L(x)|_{x_1}\;......\boldsymbol{\sigma}_L(x)|_{x_p}\right]^T, \qquad (18)$$

$$B=\left[\boldsymbol{\sigma}_L(x)f(x)|_{x_1}\;......\boldsymbol{\sigma}_L(x)f(x)|_{x_p}\right]^T, \qquad (19)$$

$$C=\begin{bmatrix}\frac{1}{4}\left(\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1}g^T(x)\nabla\boldsymbol{\sigma}_L^T\right)|_{x_1}\;...... \\ \frac{1}{4}\left(\nabla\boldsymbol{\sigma}_Lg(x)R^{-1}g^T(x)\nabla\boldsymbol{\sigma}_L^T\right)|_{x_p}\end{bmatrix}^T, \qquad (20)$$

$$D=\left[Q(x)|_{x_1}\;...Q(x)|_{x_p}\right]^T, \qquad (21)$$

where p in x_p represents the number of points of the mesh. Reducing the mesh size, we have

$$\begin{aligned}&\left\langle -\dot{\mathbf{w}}_L^T(t)\boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega \\ &=\lim_{\|\Delta x\|\to 0}-\left(A^TA\right)\cdot\dot{\mathbf{w}}_L(t)\cdot\Delta x\end{aligned}, \qquad (22)$$

$$\begin{aligned}&\left\langle -\mathbf{w}_L^T(t)\nabla\boldsymbol{\sigma}_L(x)f(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega \\ &=\lim_{\|\Delta x\|\to 0}-\left(A^TB\right)\cdot\mathbf{w}_L(t)\cdot\Delta x\end{aligned}, \qquad (23)$$

$$\begin{aligned}&\left\langle \begin{matrix}\frac{1}{4}\mathbf{w}_L^T(t)\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1} \\ \cdot g^T(x)\nabla\boldsymbol{\sigma}_L^T\mathbf{w}_L(t),\boldsymbol{\sigma}_L(x)\end{matrix}\right\rangle_\Omega, \\ &=\lim_{\|\Delta x\|\to 0}A^T\mathbf{w}_L^T(t)C\mathbf{w}_L(t)\cdot\Delta x\end{aligned} \qquad (24)$$

$$\left\langle -Q(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega=\lim_{\|\Delta x\|\to 0}-\left(A^T\cdot D\right)\cdot\Delta x\,. \qquad (25)$$

This implies that (16) can be converted to

$$\begin{aligned}&\dot{\mathbf{w}}_L(t)=-\left(A^TA\right)^{-1}\mathbf{w}_L(t)A^TB \\ &+\left(A^TA\right)^{-1}A^T\mathbf{w}_L^T(t)C\mathbf{w}_L(t)-\left(A^TA\right)^{-1}A^TD\end{aligned}. \qquad (26)$$

This is a nonlinear ODE that can easily be integrated backwards using final condition $\mathbf{w}_L(t_f)$ to find the least-squares optimal NN weights. Then, the nearly optimal value function is given by

$$V_L(x,t)=\mathbf{w}_L^T(t)\boldsymbol{\sigma}_L(x),$$

and the nearly optimal control by

$$u_L(t)=-\frac{1}{2}R^{-1}g^T(x)\nabla\boldsymbol{\sigma}_L^T(x)\mathbf{w}_L(t). \qquad (27)$$

Note that in practice, we use a numerically efficient least-squares relative to solve (26) without matrix inversion.

IV. ILLUSTRATIVE EXAMPLE

We now show the power of our NN control technique for finding nearly optimal fixed-final time controllers. Consider the following linear system

$$\begin{aligned} \dot{x}_1 &= 2x_1 + 3x_2 + u_1 \\ \dot{x}_2 &= 5x_1 + 6x_2 + 2u_2 \end{aligned} \quad . \tag{28}$$

Define performance index

$$V(x(t_0), t_0) = \frac{1}{2} x(t_f)^T S(t_f) x(t_f) + \frac{1}{2} \int_{t_0}^{T} \left(x^T Q x + u^T R u \right) dt \quad .$$

Here Q and R are chosen as identity matrices. The steady-state solution of the Riccati equation can be obtained by solving the algebraic Riccati equation (ARE). The result is $\begin{bmatrix} 3.1610 & 2.8234 \\ 2.8234 & 3.6777 \end{bmatrix}$. Our algorithm should give the same steady-state value.

To find a nearly optimal time-varying controller, the following smooth function is used to approximate the value function of the system

$$V(x_1, x_2) = w_1 x_1^2 + w_2 x_1 x_2 + w_3 x_2^2 \, .$$

This is a NN with polynomial activation functions, and hence $V(0) = 0$.

Note that if $V = x^T P x$, then $P = \begin{bmatrix} w_1 & w_2/2 \\ w_2/2 & w_3 \end{bmatrix}$.

In this example, three neurons are chosen and $\mathbf{w}_L(t_f) = [10, 10, 0]$. Our algorithm was used to determine the nearly optimal time-varying control law by backwards integrating to solve (26). A least-square algorithm was used to compute $\dot{\mathbf{w}}_L(t)$ at each integration time. From figure 1 it is obvious that about six seconds from t_f, the weights converge to the solution of the algebraic Riccati equation.

The control signal is

$$u = -\frac{1}{2} R^{-1} g^T P x \, . \tag{29}$$

The states and control signal are shown in Figures 2 and 3.

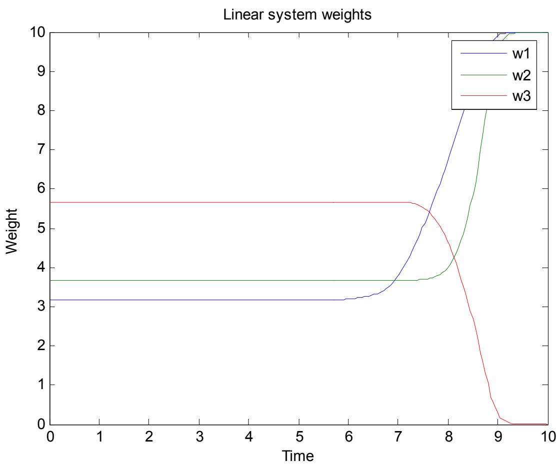

Fig. 1. Linear System Weights

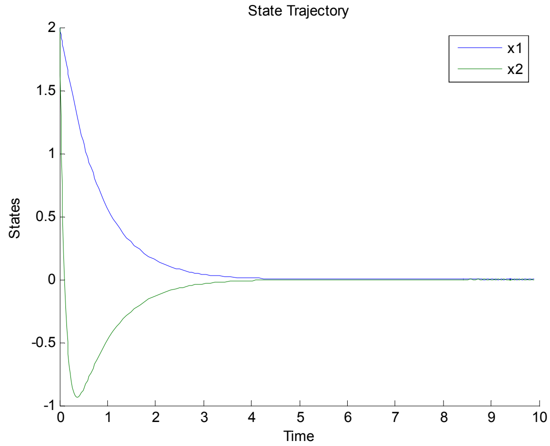

Fig. 2: State Trajectory of Linear System

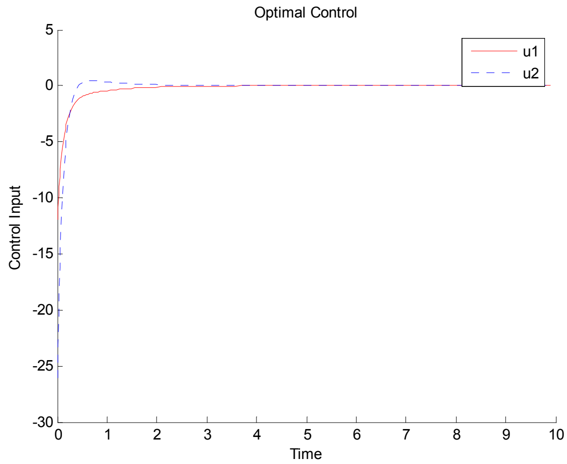

Fig. 3: Optimal NN Control Law

V. CONCLUSION

We use NN to approximately solve the time-varying HJB equation. The technique can be applied to both linear and nonlinear systems. Full conditions for convergence have been derived. Simulation examples have been carried out to show the effectiveness of the proposed method.

REFERENCES

[1] Abu-Khalaf, M., Lewis, F. L., "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach", *Automatica*, 41, 2005.

[2] Adams, R., Fournier, J., "Sobolev spaces", 2nd ed., New York: Academic Press, 2003.

[3] Balas, M. J., "Modal control of certain flexible dynamic systems", *SIAM J. Control and Optimization*, Vol 16, No. 3, pp. 450-462,1978.

[4] Beard, R., *Improving the Closed-Loop Performance of Nonlinear Systems*, PhD thesis, Rensselaer Polytechnic Institute, Troy, NY 12180, 1995.

[5] Beard, R., Saridis, G., Wen, J., "Galerkin Approximations of the Generalized Hamilton-Jacobi-Bellman Equation," *Automatica*, Vol. 33, December, pp. 2159-2177,1997.

[6] Bernstein, D. S., "Optimal nonlinear, but continuous, feedback control of systems with saturating actuators," *International Journal of Control*, Vol 62, No. 5, pp. 1209-1216.

[7] Bertsekas, D. P., Tsitsiklis, J. N., "Neuro-Dynamic Programming: An Overview," *Proc. IEEE Conference on. Decision and Control*, pp.

560-564, December 1995

[8] Chen, F. C. and Liu, C. C., “Adaptively controlling nonlinear continuous-time systems using multiplayer neural networks”, *IEEE Trans. Automat. Control*, vol. 39, No. 6, pp. 1306-1310, June 1994.

[9] Cheng, T., Frank, L. L., Abu-Khalaf, M., “A Neural Network Solution for Fixed-Final Time Optimal Control of Nonlinear Systems”, Submitted to *Automatica*.

[10] Finlayson, B. A., *The method of weighted residuals and variational principles*, New York: Academic Press, 1972

[11] Ge, S. S., “Robust adaptive NN feedback linearization control of nonlinear systems”, *International Journal of Systems Science*, Vol. 27, No. 12, pp. 1327-1338, 1996.

[12] Hornik, K., Stinchcombe, M., White, H., “Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, 3, 551-560, 1990.

[13] Huang, J., Lin, C. F., “Numerical approach to computing nonlinear H_∞ control laws”, *Journal of Guidance, Control, and Dynamics*, 18 (5), 989-994, 1995.

[14] Kim, Y. H., Lewis, F. L., Dawson. D., “Intelligent optimal control of robotic manipulators using neural networks,” *Automatica* 36, pp. 1355-1364, 2000.

[15] Lewis, F. L., Jagannathan, S., Yesildire, A., *Neural Network Control of Robot Manipulators and nonlinear Systems*, Taylor & Francis, 1999.

[16] Lewis, F. L., Syrmos, V. L., *Optimal Control*, John Wiley & Sons, Inc. New York, NY, 1995.

[17] Liu, X., Balakrishnan, S. N., “Adaptive Critic Based Neuro-Observer,” *Proc. American Control Conference*, pp. 1616-1621, June. 2001

[18] Lyshevski, S. E., *Control Systems Theory with Engineering Applications*, Birkhauser, Boston, MA, 2001.

[19] Lyshevski, S. E., “Optimal Control of Nonlinear Continuous-Time Systems: Design of Bounded Controllers Via Generalized Nonquadratic Functionals,” *Proc. American Control Conference*, pp. 205-209, June. 1998.

[20] Lyshevski, S. E., Meyer, A. U., “Control System Analysis and Design Upon the Lyapunov Method,” *Proc. American Control Conference*, pp. 3219-3223, June. 1995.

[21] Miller, W. T., Sutton, R., Werbos, P., *Neural Networks for Control*, The MIT Press, Cambridge, Massachusetts, 1990.

[22] Munos, R., Baird, L. C., Moore, A., “Gradient Descent Approaches to Neural-Net-Based Solutions of the Hamilton-Jacobi-Bellman Equation,” *International Joint Conference on Neural Networks*, Vol. 3, pp. 2152-2157, 1999.

[23] Parisini, T., Zoppoli, R. “Neural Approximations for Infinite-Horizon Optimal Control of Nonlinear Stochastic Systems”, *IEEE Trans Neural Net*. Vol. 9, No. 6, pp. 1388-1408, November 1998.

[24] Polycarpou, M. M., “Stable adaptive neural control scheme for nonlinear systems”, *IEEE Trans. Automat. Control*, vol. 41, No. 3, pp. 447-451, Mar 1996.

[25] Rovithakis, G. A. and Christodoulou, M. A., “Adaptive control of unknown plants using dynamical neural networks”, *IEEE Trans. Systems, Man, and Cybernetics*, vol. 24, No. 3, pp. 400-412, 1994.

[26] Sanner, R. M. and Slotine, J. J. E., “Stable adaptive control and recursive identification using radial gaussian networks,” *Proc. IEEE Conf. Decision and Control*, pp. 2116-2123, Brighton, 1991.

[27] Saridis, G., Lee, C. S., “An Approximation Theory of optimal Control for Trainable Manipulators”, *IEEE Trans. Systems, Man, Cybernetics*, Vol. 9, No. 3, pp. 152-159, March 1979.