Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Neural Network -based Nearly Optimal Hamilton-Jacobi-Bellman Solution for Affine Nonlinear Discrete-Time Systems[2]

Z. Chen and S. Jagannathan[1]

Abstract: **In this paper, we consider the use of nonlinear networks towards obtaining nearly optimal solutions to the control of nonlinear discrete-time systems. The method is based on least-squares successive approximation solution of the Generalized Hamilton-Jacobi-Bellman (HJB) equation. Since successive approximation using the GHJB has not been applied for nonlinear discrete-time systems, the proposed recursive method solves the GHJB equation in discrete-time on a well-defined region of attraction. The definition of GHJB, Pre-Hamiltonian function, HJB equation and method of updating the control function for the affine nonlinear discrete time systems are proposed. A neural network is used to approximate the GHJB solution. It is shown that the result is a closed-loop control based on a neural network that has been tuned a priori in off-line mode. Numerical example show that for nonlinear discrete-time systems, the updated control laws will converge to the suboptimal control.**

I. INTRODUCTION

In the literature, there are many methods of designing stable control of nonlinear systems. Ensuring optimality guarantees the stability of the nonlinear system; however, optimal control of nonlinear systems is a difficult and challenging area. If the system is modeled by linear dynamics and the cost functional to be minimized is quadratic in the state and control, then the optimal control is a linear feedback of the states, where the gains are obtained by solving a standard Riccati equation [6]. On the other hand, if the system is modeled by the nonlinear dynamics or the cost functional is non quadratic, the optimal state feedback control will depend upon obtaining the solution to the Hamilton-Jacobi-Bellman (HJB) [14] equation which is generally nonlinear. The HJB equation is difficult to solve directly because it involves solving either nonlinear partial difference or differential equations.

To overcome the difficulty in solving the HJB equation, recursive methods iteratively solve the generalized HJB (GHJB) equation, which is linear in the cost function of the system, and then update the control law. It is demonstrated [7] in the literature that if the initial control is admissible and the GHJB equation can be solved exactly, the updated control will converge to the optimal control, which is the unique solution to the HJB equation.

There has been a great deal of effort to solve the HJB equation in the literature both in continuous and discrete-time. Approximate HJB solution has been confronted using many techniques by Saridis [7], Beard [13][14], Bertsekas and Tsitsiklis[1], Lewis [5][6], and others. Since neural networks (NNs) can effectively extend adaptive control techniques to nonlinearly parameterized systems, Werbos [11] first proposed NN-based optimal control laws using the HJB equation. Although many papers, [5][7][11-14], have discussed the GHJB method for continuous-time systems, there is minimal work done for discrete-time nonlinear systems.

Discrete-time version of the approximate GHJB equation-based control is important since all the controllers are typically implemented using embedded digital hardware. In this paper, we will apply the idea of GHJB equation in discrete-time and set up the practical method for obtaining the nearly optimal control of nonlinear discrete-time systems. We use successive approximation techniques in the least-squares sense to solve the GHJB in discrete-time using a quadratic functional. A NN is used to approximate the GHJB. It is shown that the result is a closed-loop control based on a NN that has been tuned a priori in off-line mode.

II. OPTIMAL CONTROL AND GENERALIZED HAMILTON-JACOBI-BELLMAN EQUATION IN DISCRETE-TIME

Consider an affine in the control nonlinear discrete-time dynamic system of the form

$$x(k+1) = f(x(k)) + g(x(k))u(x(k)) \quad (1)$$

where $x(k) \in \Omega \subset \Re^n$, $u : \Re^n \to \Re^m$, $f : \Re^n \to \Re^n$ and $g : \Re^m \to \Re^n$. Assume that $f + gu$ is Lipschitz continuous on a set Ω in $\Re^n$ contain the origin, and that the system (1) is controllable in the sense that there exists a continuous control on Ω that asymptotically stabilizes the system. It is desired to find a control function $u : \Re^n \to \Re^m$, which minimizes the generalized quadratic cost functional

$$J(u; x(0)) = \sum_{k=0}^{N-1} \left(Q(x(k)) + u(x(k))^T Ru(x(k)) \right) + \phi(x(N)) \quad (2)$$

[1] Authors are with Department of Electrical and Computer Engineering, University of Missouri-Rolla, Rolla, MO 65409. Email:sarangap@umr.edu.

[2] Research supported in part by NSF grants ECS #0296191 and ECS #0327877.

where $Q:\Re^n \to \Re$ is a semi definite monotonically increasing function of state $x(k)$, R is a constant symmetric positive definite matrix, $\phi:\Re^n \to \Re$ is final state punishment function that is positive definite.

Control Objective: The objective is to select the feedback control law $u(k)$ to minimize the cost function.

Remark 1: It is important to note that the control u should stabilize the system on Ω and make the cost functional value finite so that the control is admissible [14].

Definition 2.1 (**Admissible Controls):** Let $\psi(\Omega)$ denotes the set of admissible controls. A control function $u:R^n \to R^m$ is defined to be admissible with respect to the state penalty function $Q(x(k))$ on a controlled invariant set Ω, denoted as $u \in \Psi(\Omega)$, if:

- u is continuous on Ω,
- $u(0)=0$,
- u stabilizes system (1) on Ω,
- $\sum_{k=0}^{\infty}\left(Q(x(k))+u(x(k))^T Ru(x(k))\right)+\phi(x(\infty))<\infty \ \forall x(k)\in\Omega$

Remark 2: The admissible control guarantees that the control is stable but not vice versa [14]. Therefore, we should restrict the systems that decay sufficiently fast.

Given an admissible control and the state of the system, the performance of this control is evaluated through a cost functional. A suitable cost function, which is independent of the solution of the nonlinear dynamic system $x(k)$, is necessary. Theorem 1 will prove that there exists a positive definite function, $V(x)$, referred as value function, whose initial value, $V(x(0))$, is equal to the cost functional value of J given an initial admissible control and the state of the system.

Lemma 2.1: If $u_1(k)$ is admissible, there exists a positive definite function $V(x)$ so that it satisfies (3) and (4). ■

Theorem 2.1: Assume $u_1(k)\in\Omega_u$ is an admissible control law arbitrarily selected. If there exists a positive definite continuously differentiable value function $V(x)$ on $\Re^n$ satisfying the following

$$\frac{\partial V(x)^T}{\partial x}\left(f(x(k))+g(x(k))u(x(k))-x(k)\right)+Q(x(k))+u(x(k))^T Ru(x(k))=0 \quad (3)$$

$$V(x(N))=\phi(x(N)) \quad (4)$$

then $V(x(j),j/u_1)$ is the value function of the system defined in (1) for all $j=0,....N$ and

$$V(x(0),u_1)=J(u_1;x(0)) \quad (5)$$

■

For convenience, denote

$$\frac{\partial V(k)^T}{\partial x(k)}=\frac{\partial V(x(k))^T}{\partial x(k)},\ f(k)=f(x(k)),\ g(k)=g(x(k)),\ u(k)=u(x(k)),$$
$$Q(k)=Q(x(k)),\ V(x(N),N/u_1)=V(x(N)) \quad (6)$$

Remark 3: An optimal control function for a nonlinear discrete-time system is the one that uses the minimum value function $V(x(0),u_1)$ and minimizes the cost functional $J(u_1;x(0))$.

Definition 2.2 (GHJB Equation in Discrete Time): The GHJB equation can be defined as

$$\frac{\partial V(k)^T}{\partial x(k)}\left(f(k)+g(k)u(k)-x(k)\right)+Q(k)+u(k)^T Ru(k)=0 \quad (7)$$

$$V(0)=0 \quad (8)$$

In this paper, the infinite-time optimal control problem for the nonlinear discrete-time system (1) is attempted. The cost functional of the infinite-time problem for the discrete time system is defined as

$$J(u;x(0))=\sum_{k=0}^{\infty}\left(Q(k)+u(k)^T Ru(k)\right) \quad (9)$$

The GHJB equation (7) with the boundary condition (8) can be used for the infinite-time regulator problems, because as $N\to\infty$, $x(\infty)=0$, $V(x(\infty))=\phi(x(\infty))=0$. So if an admissible control is specified, for any infinite-time problem, we can solve the GHJB equation to obtain the value function $V(x)$ which in turn can be used in the cost functional, J, along with $V(x(0))$ to calculate the cost of the admissible control.

Our objective is to improve the performance of the system over time by minimizing a certain pre-Hamiltonian function so that a near optimal controller results. Next, the pre-Hamiltonian function for the discrete-time system is introduced.

Definition 2.3 (Pre-Hamiltonian Function): A suitable pre-Hamiltonian function for the nonlinear system (1) is defined by

$$H(V(k),x(k),u,)=\frac{\partial V(k)^T}{\partial x(k)}\left(f(k)+g(k)u(k)-x(k)\right)+Q(k)+u(k)^T Ru(k) \quad (10)$$

It is important to note that the pre-Hamiltonian is a nonlinear function of the state, value and the control functions. If a control function $u^{(i)}\in\Omega_u$ and value function, $V^{(i)}$, satisfies $GHJB\left(V^{(i)},u^{(i)}\right)=0$, an updated control function u^{i+1} can be obtained by differentiating the pre-Hamiltonian function (10) associated with the value function $V^{(i)}$. In other words, the updated control function can be obtained by solving

$$\frac{\partial H\left(x,V^{(i)}(x),u^{(i+1)}\right)}{\partial u^{(i+1)}}=0 \quad (11)$$

and it is given by

$$u^{(i+1)} = -\frac{1}{2} R^{-1} g^T(k) \frac{\partial V^{(i)}(k)}{\partial x(k)} \qquad (12)$$

The next theorem demonstrates that the updated control function is indeed admissible for the nonlinear discrete-time system described by (1).

Theorem 2.2 (Improved Control): If $u^{(i)} \in \Psi(\Omega)$, and $V^{(i)}$ satisfies $GHJB(V^{(i)}, u^{(i)}) = 0$ with the boundary condition $V^{(i)}(0) = 0$, then the updated control function derived in (20) by using the pre-Hamiltonian results in an admissible control for the system (1) on Ω. Moreover, if $V^{(i+1)}$ is the unique positive definite function satisfying $GHJB(V^{(i+1)}, u^{(i+1)}) = 0$, then

$$V^{(i+1)}(j) \le V^{(i)}(j), \; j = 0, \cdots, N. \qquad (13)$$

Proof: Admissibility: Since $V^{(i)}$ is continuously differentiable and $g : \Re^m \to \Re^n$ is a Lipschitz continuous function on the set Ω in $\Re^n$, the new control law $u^{(i+1)}$ is continuous. Since $V^{(i)}$ is positive definite function, it attains a minimum at the origin, and thus, $\frac{\partial V^{(i)}}{\partial x}$ must vanish. This implies that $u^{(i+1)}(0) = 0$.

Taking the first difference of $V^{(i)}(k)$ along the system $(f, g, u^{(i+1)})$ trajectories to obtain

$$\Delta V^{(i)}(k) = \frac{\partial V^{(i)}(k)}{\partial x(k)}^T \left(f(k) + g(k)u^{(i+1)}(k) - x(k)\right) \qquad (14)$$

Rewriting (14) as

$$\Delta V^{(i)}(k) = \frac{\partial V^{(i)}(k)}{\partial x(k)}^T (f(k) - x(k)) + \frac{\partial V^{(i)}(k)}{\partial x(k)}^T g(k)u^{(i+1)}(k) \qquad (15)$$

Rewriting the GHJB equation $GHJB(V^{(i)}, u^{(i)}) = 0$

$$\frac{\partial V^{(i)}(x)}{\partial x}^T \left(f(k) + g(k)u^{(i)}(k) - x(k)\right) + Q(k) + u^{(i)}(k)^T R u^{(i)}(k) = 0 \qquad (16)$$

Rewriting (16) to get

$$\begin{aligned} \frac{\partial V^{(i)}(k)}{\partial x(k)}^T (f(k) - x(k)) &= \\ -\frac{\partial V^{(i)}(k)}{\partial x(k)}^T g(k)u^{(i)}(k) &- u^{(i)}(k)^T R u^{(i)}(k) - Q(k) \end{aligned} \qquad (17)$$

Substituting (17) into (15), (15) can be rewritten as

$$\begin{aligned} \Delta V^{(i)}(k) &= -u^{(i)}(k)^T R u^{(i)}(k) - Q(k) \\ &+ \frac{\partial V^{(i)}(k)}{\partial x(k)}^T g(k)\left(u^{(i+1)}(k) - u^{(i)}(k)\right) \end{aligned} \qquad (18)$$

The updated control function from (12) can be expressed for convenience as

$$\frac{\partial V(k)}{\partial x(k)}^T g(k) = -2u^{(i+1)}(k)^T R \qquad (19)$$

Substituting (19) into (18), the first difference is obtained as

$$\begin{aligned} \Delta V^{(i)}(k) &= -u^{(i)}(k)^T R u^{(i)}(k) - Q(k) \\ &- 2u^{(i+1)}(k)^T R\left(u^{(i+1)}(k) - u^{(i)}(k)\right) \end{aligned} \qquad (20)$$

Rewriting (20) as

$$\begin{aligned} \Delta V^{(i)}(k) &= -Q(k) - \left(u^{(i+1)}(k) - u^{(i)}(k)\right)^T R\left(u^{(i+1)}(k) - u^{(i)}(k)\right) \\ &- u^{(i+1)}(k)^T R u^{(i+1)}(k) \end{aligned} \qquad (21)$$

Since R is $m \times m$ positive definite matrix and Q(k) is a semi definite monotonically increasing function, the first difference can be rewritten as

$$\begin{aligned} \Delta V^{(i)}(k) &= -Q(k) - \left(u^{(i+1)}(k) - u^{(i)}(k)\right)^T R\left(u^{(i+1)}(k) - u^{(i)}(k)\right) \\ &- u^{(i+1)}(k)^T R u^{(i+1)}(k) \le 0 \end{aligned} \qquad (22)$$

This implies that the first difference of $V^{(i)}(k)$ along the system $(f, g, u^{(i+1)})$ trajectories is non positive. Thus $V^{(i)}(k)$ is a Lyapunov function for $u^{(i+1)}$ on Ω. Following the Definition 1.1, one can conclude that the updated control function $u^{(i+1)}$ is admissible on Ω.

Improved control: For second part of the Theorem 2.2, the difference between $V^{(i)}(j)$ and $V^{(i+1)}(j)$ along the trajectories $(f, g, u^{(i+1)})$ can be evaluated by using (12) and $V^{(i)}(N) = V^{(i+1)}(N) = \phi(N)$ as

$$\begin{aligned} &V^{(i+1)}(j) - V^{(i)}(j) = \\ &-\sum_{k=j}^{N-1} \left(\left(\frac{\partial V^{(i+1)}(k)}{\partial x(k)} - \frac{\partial V^{(i)}(k)}{\partial x(k)} \right)^T \left(f(k) + g(k)u^{(i+1)}(k) - x(k)\right) \right) \end{aligned} \qquad (23)$$

Rewriting (23) as

$$\begin{aligned} V^{(i+1)}(j) - V^{(i)}(j) = -\sum_{k=j}^{N-1} [&\frac{\partial V^{(i+1)}(k)}{\partial x(k)}^T .(f(k) - x(k)) + \frac{\partial V^{(i+1)}(k)}{\partial x(k)}^T g(k)u^{(i+1)}(k) \\ &- \frac{\partial V^{(i)}(k)}{\partial x(k)}^T .(f(k) - x(k)) - \frac{\partial V^{(i)}(k)}{\partial x(k)}^T g(k)u^{(i+1)}(k)] \end{aligned} \qquad (24)$$

Given the GHJB equation, since $GHJB(V^{(i)}, u^{(i)}) = 0$, one can conclude that $GHJB(V^{(i+1)}, u^{(i+1)}) = 0$. Therefore

$$\begin{aligned} \frac{\partial V^{(i)}(k)}{\partial x(k)}^T (f(k) - x(k)) &= \\ -\frac{\partial V^{(i)}(k)}{\partial x(k)}^T g(k)u^{(i)}(k) &- Q(k) - u(k)^{(i)^T} R u^{(i)}(k) \end{aligned} \qquad (25)$$

and

$$\begin{aligned} \frac{\partial V^{(i+1)}(k)}{\partial x(k)}^T (f(k) - x(k)) &= \\ -\frac{\partial V^{(i+1)}(k)}{\partial x(k)}^T g(k)u^{(i+1)}(k) &- Q(k) - u(k)^{(i+1)^T} R u^{(i+1)}(k) \end{aligned} \qquad (26)$$

Adding (25) and (26) into (24) we get

$$V^{(i+1)}(j) - V^{(i)}(j) = \sum_{k=j}^{N-1} [u^{(i+1)}(k)^T R u^{(i+1)}(k) - \frac{\partial V^{(i)}(k)}{\partial x}^T g(k)u^{(i)}(k)$$

$$-u^{(i)}(k)^T Ru^{(i)}(k)+\frac{\partial V^{(i)}(k)^T}{\partial x} g(k)u^{(i+1)}(k)] \quad (27)$$

Using (19) into (27), (27) can be expressed as

$$V^{(i+1)}(j)-V^{(i)}(j)=\sum_{k=j}^{N-1}[u^{(i+1)}(k)^T Ru^{(i+1)}(k)+2u^{(i+1)}(k)^T Ru^{(i)}(k)$$
$$-u^{(i)}(k)^T Ru^{(i)}(k)-2u^{(i+1)}(k)^T \cdot R.u^{(i+1)}(k)] \quad (28)$$

With further simplification, we get

$$V^{(i+1)}(j)-V^{(i)}(j)= -\sum_{k=j}^{N-1}\left(\left(u^{(i+1)}(k)-u^{(i)}(k)\right)^T R\left(u^{(i+1)}(k)-u^{(i)}(k)\right)\right) \quad (29)$$

Since R is a $m\times m$ positive definite matrix, it can be concluded that

$$V^{(i+1)}(j)-V^{(i)}(j)= -\sum_{k=j}^{N-1}\left(\left(u^{(i+1)}(k)-u^{(i)}(k)\right)^T R\left(u^{(i+1)}(k)-u^{(i)}(k)\right)\right)\le 0 \quad (30)$$

Thus

$$V^{(i+1)}(j)\le V^{(i)}(j), j=0,\cdots,N.$$

Theorem 2.2 suggests that after solving the GHJB equation and updating the control function by using (12) over time the updated control will converge close to the solution of HJB, which then renders the optimal control function. The GHJB becomes the Hamilton Jacobi-Bellman (HJB) equation on substitution of the optimal control function $u^*(x)$. The HJB equation can be defined as follows. ■

Definition 2.4 (**HJB Equation in Discrete Time):** The HJB equation in discrete-time can be expressed as

$$\frac{\partial V^*(k)^T}{\partial x}(f(k)-x(k))+Q(k)-\frac{1}{4}\frac{\partial V^*(k)^T}{\partial x} gR^{-1}g^T(k)\frac{\partial V^*(k)}{\partial x(k)}=0 \quad (31)$$

$$V^*(0)=0 \quad (32)$$

where the optimal control function is given by

$$u^*=-\frac{1}{2}R^{-1}g^T(k)\frac{\partial V^*(k)}{\partial x(k)} \quad (33)$$

Note V^* is the unique optimal solution to the HJB equation (31). Note that the GHJB is linear in the value function derivative while the HJB equation is nonlinear in the value function derivative. Solving the GHJB equation requires solving linear partial difference equations, while the HJB equation solution involves nonlinear partial difference equations, which may be difficult to solve. This is the reason for introducing the successive approximation technique using GHJB. In the successive approximation method, one solves (7) for $V(x)$ given a stabilizing control $u(k)$ then finds an improved control based on $V(k)$ using (12). In the following, Corollary 1 indicates that if the initial control function is admissible, then repetitive application of (12), the sequence of solutions $V^{(i)}$ converges to the optimal HJB solution $V^*(k)$.

Corollary 2.1 (Convergence of Successive Approximations): Given an initial admissible control, $u^0(x)\in\Psi(\Omega)$, by iteratively solving GHJB equation and updating the control function using (12), the sequence of solutions $V^i(x)$ will converge to the optimal HJB solution $V^*(x)$.

III. Neural Network Least-Squares Approach

The purpose of this section is to show how we approximate the solution of the GHJB equation in discrete-time using NNs such that the controls which result from the solution are in feedback form. Using the NN approximation property [4] in compact set we approximate $V(x)$ with a NN

$$V_L(x)=\sum_{j=1}^{L}w_j\sigma_j(x)=W_L^T\bar{\sigma}_L(x) \quad (34)$$

where the activation function vector $\sigma_j(x):\Omega\to\Re$, is continuous, $\sigma_j(0)=0$ and the neural network weights are w_j and L is the number of hidden layer neurons. The vectors $\bar{\sigma}_L(x)\equiv[\sigma_1(x),\sigma_2(x),\cdots,\sigma_L(x)]^T$ and $W_L\equiv[w_1,w_2,\cdots,w_L]^T$ are the vector of activation function and NN weight matrix respectively. The NN weights will be tuned to minimize the residual error in a least-squares sense over a set of points within the stability region of the initial stabilizing control. Least squares solution [3] attains the lowest possible residual error with respect to the NN weights.

For the $GHJB(V,u)=0$, V is replaced by V_L having a residual error as

$$GHJB\left(V_L=\sum_{j=1}^{L}w_j\sigma_j,u\right)=e_L(x) \quad (35)$$

To find the least-squares solution, the method of weighted residuals is used [3]. The weights w_j are determined by projecting the residual error onto $\frac{\partial(e_L(x))}{\partial W_L}$ and setting the result to zero $\forall x\in\Omega$, i.e.

$$\left\langle\frac{\partial(e_L(x))}{\partial W_L},e_L(x)\right\rangle=0 \quad (36)$$

When expanded, the above equation becomes

$$\langle\nabla\bar{\sigma}_L(f+gu-x),\nabla\bar{\sigma}_L(f+gu-x)\rangle W_L + \left\langle Q+u^TRu,\nabla\bar{\sigma}_L(f+gu-x)\right\rangle=0 \quad (37)$$

where $\nabla\bar{\sigma}_L=\left[\frac{\partial\sigma_1(x)}{\partial x},\frac{\partial\sigma_2(x)}{\partial x},\cdots,\frac{\partial\sigma_L(x)}{\partial x}\right]^T$. In order to proceed, the following technical results are needed.

Lemma 3.1: if the set $\{\sigma_j(x)\}_1^L$ is linearly independent and $u \in \psi(\Omega)$, then the set

$$\left\{ \frac{\partial \sigma_j}{\partial x}^T (f + gu - x) \right\}_1^L \quad (38)$$

is also linearly independent.

Lemma 3.2: Riemann Approximation of Integrals

An integral can be approximated as

$$\int_a^b f(x)dx = \lim_{\|\Delta x\| \to 0} \sum_{i=1}^{n} f(\bar{x}_i) \cdot \Delta x \quad (39)$$

where $\Delta x = x_i - x_{i-1}$ and f is bounded on $[a, b]$, [2].

Introducing a mesh on Ω, with mesh size equal to Δx, which is taken very small, we can rewrite some terms in (40) as follows:

$$X = \left[\nabla \bar{\sigma}_L (f + gu - x)|_{x1} \cdots \nabla \bar{\sigma}_L (f + gu - x)|_{x_p} \right]^T \quad (40)$$

$$Y = \begin{bmatrix} Q + u^T Ru \ |_{x_1} \\ \vdots \\ Q + u^T Ru \ |_{x_p} \end{bmatrix} \quad (41)$$

where p in x_p represents the number of points of the mesh. This number increases as the mesh size is reduced.

Using Lemma 3.2, we can rewrite (40) as

$$XW_L + Y = 0 \quad (42)$$

This implies that we can calculate

$$W_L = -(X^T X)^{-1}(XY) \quad (43)$$

An interesting observation is that equation (43) is the standard least-squares method of estimation for a mesh on Ω. Note that the mesh size Δx should be such that the number of points p is greater or equal to the order of the approximation L and the activation functions should be linearly independent. These conditions guarantee a full rank for $(X^T X)$.

The optimal control of nonlinear discrete-time system can be obtained off line by going through six steps:

Define a NN as $V = \sum_{j=1}^{L} w_j \sigma_j(x)$ to approximate smooth function of $V(x)$;

Select an admissible feedback control law u_1;

Find $V^{(1)}$ associated with u_1 to satisfy GHJB by applying least square method (LSM) to obtain the NN weights W^1;

Update the control as

$$u_2 = -\frac{1}{2} R^{-1} g(k)^T \frac{\partial V^{(1)}(k)}{\partial x(k)} \quad (44)$$

Find $V^{(2)}$ associated with u_2 to satisfy GHJB by using LSM to obtain W^2;

If $V^1(0) - V^2(0) \leq \varepsilon$, where ε is a small positive constant, then $V^* = V^{(1)}$ and stop. Otherwise, go back to step 4 by increasing the index by one. After we get V^*, the optimal state feedback control, which can be implemented online, can be described as

$$u^* = -\frac{1}{2} R^{-1} g^T(k) \frac{\partial V^*(k)}{\partial x(k)} \quad (45)$$

IV. NUMERICAL EXAMPLE

A real-world two-link planar robot arm system was used to demonstrate that the proposed approach renders a suboptimal solution for nonlinear discrete-time systems. In all of the examples, the basis functions required will be obtained from even polynomials so that the NN can approximate the positive definite function $V(x)$. If the dimension of the system is n and the order of approximation is M, then we use all of the terms in expansion of the polynomial [14]

$$\sum_{j=1}^{M/2} \left(\sum_{k=1}^{n} x_k \right)^{2j} \quad (46)$$

The resulting basis functions for a two dimensional system is given by

$$\{x_1^2, x_1 x_2, x_2^2, x_1^4, x_1^3 x_2, \cdots, x_2^M\} \quad (47)$$

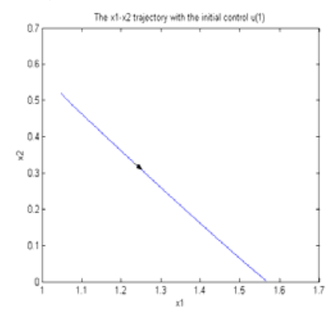

Figure 1: State trajectory with initial admissible control.

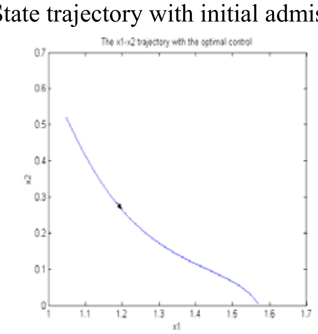

Figure 2: State trajectory with GHJB-based control.

Example (Two-link Planar Revolute-Revolute (RR) Robot Arm System): A two-link planar RR robot arm used extensively for simulation in the literature is considered. The dynamics of the two-link robot arm system is obtained by discretizing the continuous time dynamics. The control objective is moving the arm from an initial state to the final state with the cost function defined as

$$J = \int_{t=0}^{\infty} \left(\|x(t) - x_d\|^2 + \|u(t)\|^2 \right) dt \tag{48}$$

First, we will convert the continuous time dynamics system and cost function into discrete time. Let us consider a discrete-time system with a sampling period Δt and denote a time function $f(t)$ at $t = k\Delta t$ as $f(k)$, where k is a sampling number. Choosing the sampling period Δt is sufficiently small, we use the following approximation for the derivative of $f(k)$ as

$$\dot{f}(k) \cong \frac{1}{\Delta t}\left(f(k+1) - f(k)\right) \tag{49}$$

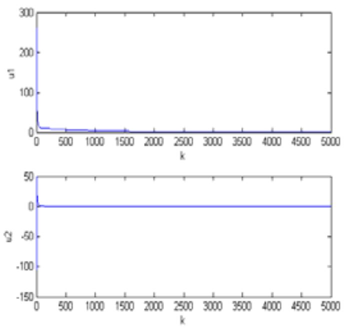

Figure 3: Control input using initial admissible control.

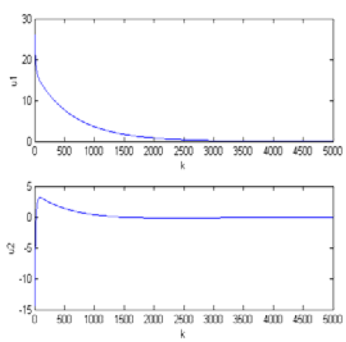

Figure 4: Control input using HJB-based control.

A sampling interval of $\Delta t = 1ms$ is used. Control function updating rule is taken as

$$u_{i+1} = -\frac{1}{2} R^{-1} g'^{T} \nabla V^{(i)}(k) \tag{50}$$

The u_i and V^i satisfy the GHJB equation:

$$\nabla V^{(i)}(x)^T \left(f'(x) + g'(x)u^{(i)}(x) - x\right) + (x - x_d)^T T (x - x_d) + u^{(i)T} R u^{(i)} = 0 \tag{51}$$

where $T = 0.001 \times I^4$ and $R = 0.001 \times I^4$.

In the simulation, the mesh size Δx is selected as 0.2, the asymptotic stability region is chosen as $0 \le x_1 \le 2$, $-1 \le x_2 \le 1$, $-1 \le x_3 \le 1$, $-1 \le x_4 \le 1$. The small positive constant is selected as $\varepsilon = 0.01$. We use the GHJB method to obtain the nearly optimal control. After updated 5 times, the control has converged to the nearly optimal control u^*.

From the results, the trajectory with nearly optimal control (Figure 2) is a little longer than the trajectory with initial admissible control (Figure 1) even though the cost functional value (Figures 3 and 4) with optimal control is significantly less for GHJB-based control. This is due to the tradeoff observed between the trajectory selection and the control input.

V. Conclusions

In this paper**,** HJB, GHJB and pre-Hamiltonian functions in discrete-time are introduced. A systematic method of obtaining the optimal control for general affine nonlinear discrete-time system is proposed. Given an initial admissible control, the improved control through neural network successive approximation of the GHJB equation renders an optimal control for linear systems whereas for nonlinear discrete time system, the updating control law will converge to a suboptimal control.

References

[1] Bertsekas, D. P. and J. N. Tsitsiklis, *Neuro-Dynamic Programming,* Athena Scientific, Belmont, MA, 1996.

[2] Burk, F., *Lebesgue Measure and Integration,* John Wiley & Sons, New York, NY, 1998.

[3] Finlayson, B. A., *The Method of Weighted Residuals and Variational Principles.*, Academic Press, New York, NY, 1972.

[4] F. L. Lewis, S. Jagannathan and A. Yesilderek, *Neural Network Control of Robot Manipulator and Nonlinear Systems*, Taylor & Francis Inc., UK, 1999.

[5] F. L. Lewis and Murad Abu-Kuhalf, "A Hamilton-Jacobi setup for constrained neural network control", *Proc. Of the International Symposium on Intelligent Control*, pp. 1-15, 2003.

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

[7] George N. Saridis and C. S. Lee, "An approximation theory of optimal control for trainable manipulators", *IEEE Transactions on System, Man, and Cybernetics*, vol. 9, no. 3, pp. 152-159, 1979.

[8] H. Kawasaki, and Geng Li, "Gain tuning in discrete-time adaptive control for robots", *Proc. of the SICE Annual Conference in Fukui*, August 4-6, pp. 1286-1291, 2003.

[9] Kleinman, D., "On a iterative technique for Riccati equation computations", *IEEE Trans Automatic Control*, vol. 13, no. 1, pp. 114-115, 1968.

[10] Lyshevski, S. E., *Control Systems Theory with Engineering Applications,* Birkhauser, Boston, MA, 1990.

[11] Miller, W. T., R. Sutton, and P. Werbos, *Neural Networks for Control,* the MIT Press, Cambridge, Massachusetts.

[12] Ming Xin and S.N. Balakrishnan, "A new method for suboptimal control of a class of nonlinear systems", *Proceedings of the 41st IEEE Conference on Decision and Control*, vol. 3, pp. 2756-2761, 2002.

[13] R. W. Beard, George N. Saridis and John T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation", *Automatica*, vol. 33, No.12 pp. 2159 2177, 1997.

[14] R. W. Beard, *Improving the closed-loop performance of nonlinear systems,* PhD Thesis, Rensselaer Polytechnic Institute, 1995.