Available online at www.sciencedirect.com
ScienceDirect
automatica
Automatica 43 (2007) 482–490
www.elsevier.com/locate/automatica

Brief paper

A neural network solution for fixed-final time optimal control of nonlinear systems ☆

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

Automation and Robotics Research Institute, The University of Texas at Arlington, TX 76118, USA

Received 4 August 2005; received in revised form 21 July 2006; accepted 17 September 2006

Abstract

In this paper, fixed-final time optimal control laws using neural networks and HJB equations for general affine in the input nonlinear systems are proposed. The method utilizes Kronecker matrix methods along with neural network approximation over a compact set to solve a time-varying HJB 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 an example.

Keywords: Finite-horizon optimal control; Hamilton–Jacobi–Bellman; Neural networks

1. Introduction

In many practical engineering problems, one is interested in finding finite-time optimal control laws for nonlinear systems. It is known that this optimization problem (Lewis & Syrmos, 1995), requires solving a time-varying Hamilton–Jacobi–Bellman (HJB) equation that is hard to solve in most cases. Approximate HJB solutions have been found using many techniques such as those developed by Saridis and Lee (1979), Beard (1995), Bertsekas and Tsitsiklis (1995) and Kim, Lewis, and Dawson (2000). Huang and Lin (1995) provided a Taylor series expansion of the HJI equation which is closely related to the HJB equation. A local H_∞ controller is derived in Aguilar, Orlov, and Acho (2003) using perturbation methods.

Successful neural networks (NNs) controllers not based on optimal techniques have been reported in Chen and Liu (1994). 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 Miller, Sutton, and Werbos (1990). Parisini and Zoppoli (1998) 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. It is shown that using a NN approach, one can simply transform the problem into solving an ordinary differential equation (ODE) backwards in time.

We were motivated by the important results in Beard (1995). However, in contrast to that work, we are able to approximately solve the time-varying HJB equation, and do not need to perform policy iteration using the so-called GHJB equation followed by control law updates. We accomplish this by using a NN approximation for the value function which is based on a universal basis set, and by introduction of the Kronecker product to handle bilinear terms. We also demonstrate uniform convergence results over a Sobolev space.

2. Background on fixed-final time HJB optimal control

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

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

☆ This paper was not presented in any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Bernard Brogliato under the direction of Editor H.K. Khalil. Supported by NSF ECS-0501451 and ARO W91NF-05-1-0314.

* Corresponding author. Tel./fax: +1 817 272 5938.
E-mail addresses: chengtao@arri.uta.edu (T. Cheng), lewis@uta.edu (F.L. Lewis), abukhalaf@arri.uta.edu (M. Abu-Khalaf).

0005-1098/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2006.09.021

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 $f(x) + g(x)u(t)$ is Lipschitz continuous on a set $\Omega_0 \subseteq \Re^n$ containing the origin, and that system (1) is stabilizable in the sense that there exists a continuous control that asymptotically stabilizes the system on Ω_0. It is desired to find the control $u(t)$ that minimizes a generalized nonquadratic functional

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

with $Q(x)$, $W(u)$ positive definite on Ω_0, i.e. $\forall x \neq 0$, $x \in \Omega_0$, $Q(x) > 0$ and $x = 0 \Rightarrow Q(x) = 0$. A common choice for $W(u) = u^{\mathrm{T}}Ru$, 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_0, denoted by $u \in \Psi(\Omega_0)$, if u is continuous, $u(0) = 0$, u stabilizes (1) on Ω_0, and $\forall x_0 = x(t_0) \in \Omega_0$, $V(x_0, t_0)$ is finite.

An infinitesimal equivalent to (2) is (Lewis & Syrmos, 1995)

$$-\frac{\partial V(x,t)}{\partial t} = \frac{\partial V^{\mathrm{T}}(x,t)}{\partial x}(f(x)+g(x)u(t))+Q(x)+W(u). \tag{3}$$

This is a time-varying partial differential equation that yields the value $V(x,t)$ 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

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

For unconstrained control inputs, on substitution of the optimal control (Lewis & Syrmos, 1995)

$$u^*(x) = -\frac{1}{2}R^{-1}g^{\mathrm{T}}\frac{\partial V^*(x,t)}{\partial x}, \tag{5}$$

where $V^*(x,t)$ is the value function of the optimal control problem that solves the HJB equation, (3) becomes the well-known time-varying HJB equation,

$$\begin{aligned} HJB(V^*(x,t)) = {} & \frac{\partial V^*(x,t)}{\partial t} + \frac{\partial V^*(x,t)}{\partial x} f + Q(x) \\ & - \frac{1}{4}\frac{\partial V^{*\mathrm{T}}(x,t)}{\partial x} g(x)R^{-1}g^{\mathrm{T}}(x)\frac{\partial V^*(x,t)}{\partial x} = 0. \end{aligned} \tag{6}$$

This equation provides the solution to fixed-final time optimal control for general nonlinear systems. However, this equation is generally impossible to solve.

Remark 1. Optimal control problems do not necessarily have smooth or even continuous value functions, (Huang, Wang, & Teo, 2000). Lio (2000) used the theory of viscosity solutions to show that for infinite horizon optimal control problems with unbounded cost functional, under certain continuity assumptions of the dynamics, the value function is continuous on some set Ω, $V^*(x,t) \in C(\Omega)$. In this paper, all derivations are performed under the assumption of smooth solutions to (6). A similar assumption was made by Van der Schaft (1992) and Isidori and Astolfi (1992).

3. Nonlinear fixed-final time HJB solution by NN least-squares approximation

The HJB equation (6) is difficult to solve for the cost function $V(x,t)$. In this section, NNs are used to solve approximately for the value function in (6) over Ω by approximating the cost function $V(x,t)$ uniformly in t. The result is an efficient, practical, and computationally tractable solution algorithm to find nearly optimal state feedback controllers for nonlinear systems.

3.1. NN approximation of $V(x,t)$

It is well known that a NN can be used to approximate smooth time-invariant functions on prescribed compact sets (Hornik, Stinchcombe, & White, 1990). 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. In Sandberg (1998), it is shown that NNs with time-varying weights can be used to approximate uniformly continuous time-varying functions. We assume that $V(x,t)$ is smooth, and so uniformly continuous on a compact set. Therefore, one can use the following equation to approximate $V(x,t)$ for $t \in [t_0, t_f]$ on a compact set $\Omega \subset \Re^n$:

$$V_L(x) = \sum_{j=1}^{L} w_j\sigma_j(x) = w_L^{\mathrm{T}}(t)\sigma_L(x). \tag{7}$$

The set σ_j is selected to be independent. Then without loss of generality, they can be assumed to be orthonormal, i.e. select equivalent basis functions to σ_j that are also orthonormal (Bartle, 1976). The orthonormality of the set $\{\sigma_j\}_1^\infty$ on Ω implies that if a function $\psi(x) \in L_2(\Omega)$ then

$$\psi(x) = \sum_{j=1}^{\infty}\langle\psi, \sigma_j\rangle_\Omega\sigma_j(x),$$

where $\langle f, g\rangle_\Omega = \int_\Omega gf^{\mathrm{T}}\,\mathrm{d}x$.

Note that, since one requires $\partial V(x,t)/\partial t$ in (6), the NN weights are selected to be time varying. 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 (Abu-Khalaf & Lewis, 2005). 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 (Finlayson, 1972).

Note that

$$\frac{\partial V_L(x,t)}{\partial x} = \frac{\partial \boldsymbol{\sigma}_L^{\mathrm{T}}(x)}{\partial x}\mathbf{w}_L(t) \equiv \nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)\mathbf{w}_L(t), \tag{8}$$

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

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

Therefore, approximating $V(x,t)$ by $V_L(x,t)$ uniformly in t in the HJB equation (6) results in

$$\begin{aligned}&\dot{\mathbf{w}}_L^{\mathrm{T}}(t)\boldsymbol{\sigma}_L(x) + \mathbf{w}_L^{\mathrm{T}}(t)\nabla\boldsymbol{\sigma}_L(x)f(x)\\ &\quad - \tfrac{1}{4}\mathbf{w}_L^{\mathrm{T}}(t)\boldsymbol{\sigma}_L(x)g(x)R^{-1}g^{\mathrm{T}}(x)\boldsymbol{\sigma}_L^{\mathrm{T}}(x)\mathbf{w}_L(t) + Q(x)\\ &\qquad = e_L(x,t),\end{aligned} \tag{10}$$

or

$$HJB\left(V_L(x,t) = \sum_{j=1}^{L} w_j(t)\sigma_j(x)\right) = e_L(x,t), \tag{11}$$

where $e_L(x,t)$ is a residual equation error. From (10) the corresponding optimal control input is

$$u_L(x,t) = -\tfrac{1}{2}R^{-1}g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)\mathbf{w}_L(t). \tag{12}$$

To find the least-squares solution for $\mathbf{w}_L(t)$, the method of weighted residuals is used (Finlayson, 1972). The weight derivatives $\dot{\mathbf{w}}_L(t)$ are determined by projecting the residual error onto $\partial e_L(x,t)/\partial\dot{\mathbf{w}}_L(t)$ and setting the result to zero $\forall x \in \Omega$ using the inner product, i.e.

$$\left\langle \frac{\partial e_L(x,t)}{\partial \dot{\mathbf{w}}_L(t)}, e_L(x,t)\right\rangle_{\Omega} = 0. \tag{13}$$

From (11) we can get

$$\frac{\partial e_L(x,t)}{\partial \dot{\mathbf{w}}_L(t)} = \boldsymbol{\sigma}_L(x). \tag{14}$$

Therefore, one obtains

$$\begin{aligned}\dot{\mathbf{w}}_L(t) &= \langle\boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\rangle_{\Omega}^{-1}\cdot\langle\nabla\boldsymbol{\sigma}_L(x)f(x),\boldsymbol{\sigma}_L(x)\rangle_{\Omega}\cdot\mathbf{w}_L(t)\\ &\quad + \langle\boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\rangle_{\Omega}^{-1}\cdot\left\langle \begin{matrix}\frac{1}{4}\mathbf{w}_L^{\mathrm{T}}(t)\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1}\\ \cdot\, g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}\mathbf{w}_L(t),\boldsymbol{\sigma}_L(x)\end{matrix}\right\rangle_{\Omega}\\ &\quad - \langle\boldsymbol{\sigma}_L(x),\boldsymbol{\sigma}_L(x)\rangle_{\Omega}^{-1}\cdot\langle Q(x),\boldsymbol{\sigma}_L(x)\rangle_{\Omega}\end{aligned} \tag{15}$$

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.

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

3.2. Convergence of the method of least squares

In what follows, we show convergence results as L increases for the method of least squares when NN are used to uniformly approximate the cost of function in t.

Let $F(t,x)$ be piecewise continuous in t and satisfy the Lipschitz condition

$$\|F(t,x) - F(t,y)\| \leqslant K\|x-y\|,$$

$\forall x,y \in B = \{x \in R^n \,|\, \|x - x_0\| \leqslant r\}$, $\forall t \in [t_0,t_1]$, where K is constant and $\|f\|^2 = \langle f,f\rangle$. Then, there exists some $\delta > 0$ such that the state equation $\dot{x} = F(t,x)$ with $x(t_0) = x_0$ has a unique solution over $[t_0, t_0+\delta]$. (See Khalil, 2002.)

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

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

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 (Adams & Fournier, 2003), since we require to prove the convergence of both $V_L(x,t)$ and its gradient.

We now show the following convergence results.

Lemma 1. *Convergence of approximate HJB equation.*

Given $u \in \psi(\Omega)$. Let $V_L(x,t) = \sum_{j=1}^{L} w_j^{\mathrm{T}}(t)\sigma_j(x)$ satisfy $\langle HJB(V_L(x,t)),\boldsymbol{\sigma}_L(x,t)\rangle_{\Omega} = 0$ and $\langle V_L(x_f,t_f),\boldsymbol{\sigma}_L\rangle_{\Omega} = 0$, and let $V(x,t) = \sum_{j=1}^{\infty} c_j^{\mathrm{T}}(t)\sigma_j(x)$ and $\mathbf{c}_L(t) \equiv [c_1(t)c_2(t)\cdots c_L(t)]^{\mathrm{T}}$ satisfy $HJB(V(x,t)) = 0$ and $V(x(t_f),t_f) = \phi(x(t_f),t_f)$.

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 $|HJB(V_L(x,t))| \to 0$ uniformly in t on Ω as L increases.

Proof. The hypotheses implies that $HJB(V_L(x,t))$ are in $L_2(\Omega)$. Note that

$$\begin{aligned}&\langle HJB(V_L(x,t)),\sigma_j(x)\rangle_{\Omega}\\ &\quad = \sum_{k=1}^{L}\dot{w}_k^{\mathrm{T}}(t)\langle\sigma_k(x),\sigma_j(x)\rangle_{\Omega}\\ &\qquad + \sum_{k=1}^{L} w_k^{\mathrm{T}}(x)\langle\nabla\sigma_k(x)f(x),\sigma_j(x)\rangle_{\Omega}\\ &\qquad - \sum_{k=1}^{L} w_k^{\mathrm{T}}(t)\cdot\left\langle\begin{matrix}\frac{1}{4}\nabla\sigma_k(x)g(x)R^{-1}\\ \cdot\, g^{\mathrm{T}}(x)\nabla\sigma_k^{\mathrm{T}}(x),\sigma_j(x)\end{matrix}\right\rangle_{\Omega}\cdot w_k(t)\\ &\qquad + \langle Q(x),\sigma_j(x)\rangle_{\Omega}.\end{aligned} \tag{16}$$

Then

$$
\begin{aligned}
&|HJB(V_L(x,t))| \\
&= \left|\sum_{j=1}^{\infty}\langle HJB(V_L(x,t)),\sigma_j(x)\rangle_\Omega\sigma_j(x)\right| \\
&= \left|\begin{array}{l}
\sum_{j=L+1}^{\infty}\left(\sum_{k=1}^{L}\dot{w}_k^{\mathrm{T}}(t)\langle\sigma_k(x),\sigma_j(x)\rangle_\Omega\right)\sigma_j(x) \\
+\sum_{j=L+1}^{\infty}\left(\sum_{k=1}^{L}w_k^{\mathrm{T}}(t)\langle\nabla\sigma_k(x)f(x),\sigma_j(x)\rangle_\Omega\right)\sigma_j(x) \\
-\sum_{j=L+1}^{\infty}\left(\sum_{k=1}^{L}w_k^{\mathrm{T}}(t)\left\langle\begin{array}{l}\frac{1}{4}\nabla\sigma_k(x)g(x)R^{-1}\\ \cdot g^{\mathrm{T}}(x)\nabla\sigma_k^{\mathrm{T}}(x),\sigma_j(x)\end{array}\right\rangle_\Omega w_k(t)\right)\sigma_j(x) \\
+\sum_{j=L+1}^{\infty}\langle Q(x),\sigma_j(x)\rangle_\Omega\sigma_j(x).
\end{array}\right|
\end{aligned}
\tag{17}
$$

Since the set $\{\sigma_j(x)\}_1^\infty$ are orthogonal, $\langle\sigma_k(x),\sigma_j(x)\rangle_\Omega=0$.

Therefore

$$
\begin{aligned}
&|HJB(V_L(x,t))| \\
&\leqslant \left|\left(\sum_{k=1}^{L}w_k(t)\sum_{j=L+1}^{\infty}\langle\nabla\sigma_k(x)f(x),\sigma_j(x)\rangle_\Omega\cdot\sigma_j(x)\right)\right| \\
&\quad+\left|\left(\sum_{k=1}^{L}w_k^2(t)\cdot\sum_{j=L+1}^{\infty}\left\langle\begin{array}{l}\frac{1}{4}\nabla\sigma_k(x)g(x)R^{-1}\\ \cdot g^{\mathrm{T}}(x)\nabla\sigma_k^{\mathrm{T}}(x),\sigma_j(x)\end{array}\right\rangle_\Omega\cdot\sigma_j(x)\right)\right| \\
&\quad+\left|\sum_{j=L+1}^{\infty}\langle Q(x),\sigma_j(x)\rangle_\Omega\sigma_j(x)\right| \\
&\quad\times\left|\sum_{j=L+1}^{\infty}\left(\sum_{k=1}^{L}w_k^{\mathrm{T}}(t)\langle\nabla\sigma_k(x)f(x),\sigma_j(x)\rangle_\Omega\right)\cdot\sigma_j(x)\right| \\
&\leqslant AB(x)+CD(x) \\
&\quad+Vec\left(\left|\sum_{j=L+1}^{\infty}\langle Q(x),\sigma_j(x)\rangle_\Omega\sigma_j(x)\right|\right),
\end{aligned}
\tag{18}
$$

where

$$A=\max_{1\leqslant k\leqslant L}|w_k(t)|,$$

$$B(x)=\sup_{(t,x)\in[t_0,T]\times\Omega}\left|\sum_{k=1}^{L}\left(\sum_{j=L+1}^{\infty}\langle\nabla\sigma_k(x)f(x),\sigma_j(x)\rangle_\Omega\right)\sigma_j(x)\right|,$$

$$C=\max_{1\leqslant k\leqslant L}|w_k^2(t)|,$$

$$D=\sup_{(t,x)\in[t_0,T]\times\Omega}\left|\left(\sum_{k=1}^{L}\left(\sum_{j=L+1}^{\infty}\left\langle\begin{array}{l}\frac{1}{4}\nabla\sigma_k(x)g(x)R^{-1}\\ \cdot g^{\mathrm{T}}(x)\nabla\sigma_k^{\mathrm{T}}(x),\sigma_j(x)\end{array}\right\rangle_\Omega\right)\sigma_j(x)\right)\right|.$$

Suppose $\nabla\sigma_k(x)f(x)$, $\frac{1}{4}\nabla\sigma_k(x)g(x)R^{-1}g^{\mathrm{T}}(x)\nabla\sigma_k^{\mathrm{T}}(x)$ and $Q(x)$ are in $L_2(\Omega)$, the orthonormality of the set $\{\sigma_j(x)\}_1^\infty$ implies that $B(x)$ and the second and third term on the right-hand side can be made arbitrarily small by an appropriate choice of L.

Therefore

$$A\cdot B(x)+C\cdot D(x)\to 0$$

and

$$\left|\sum_{j=L+1}^{\infty}\langle Q(x),\sigma_j(x)\rangle_\Omega\sigma_j\right|\to 0.$$

So $|HJB(V_L(x,t))|\to 0$ uniformly in t on Ω as L increases. □

Lemma 2. *Convergence of NN weights.*

Given $u\in\Psi(\Omega)$ and suppose the hypotheses of Lemma 1 hold. Then $\|\mathbf{w}_L(t)-\mathbf{c}_L(t)\|_2\to 0$ uniformly in t as L increases.

Proof. Define $e_L(x,t)=HJB(V_L(x,t))$ and

$$\hat{e}_L(x,t)=V_L(x(t_f),t_f)-\phi(x(t_f),t_f). \tag{19}$$

Then $\langle e_L(x,t),\boldsymbol{\sigma}_L(x)\rangle_\Omega=\langle\hat{e}_L(x,t),\boldsymbol{\sigma}_L(x)\rangle_\Omega=0$. From the hypotheses we have that

$$
\begin{aligned}
&HJB(V_L(x,t))-HJB(V(x,t))=e_L(x,t),\\
&(V_L-V)(x,t_f)=\hat{e}_L(x,t),
\end{aligned}
\tag{20}
$$

substituting the series expansion for $V_L(x,t)$ and $V(x,t)$, and moving the terms in the series that are greater than L to the right-hand side we obtain

$$
\begin{aligned}
&(\dot{\mathbf{w}}_L(t)-\dot{\mathbf{c}}_L(t))^{\mathrm{T}}\boldsymbol{\sigma}_L(x)+(\mathbf{w}_L(t)-\mathbf{c}_L(t))^{\mathrm{T}}\nabla\boldsymbol{\sigma}_L(x)f(x)\\
&\quad-(\mathbf{w}_L^{\mathrm{T}}(t)\otimes\mathbf{w}_L^{\mathrm{T}}(t)-\mathbf{c}_L^{\mathrm{T}}(t)\otimes\mathbf{c}_L^{\mathrm{T}}(t))\\
&\quad\cdot Vec\left(\frac{1}{4}\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1}\cdot g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)\right)\\
&\qquad=e_L(x,t)+\sum_{j=L+1}^{\infty}\dot{c}_j^{\mathrm{T}}(t)\sigma_j(x)+\sum_{j=L+1}^{\infty}c_j^{\mathrm{T}}(t)\nabla\sigma_j(x)f(x)\\
&\qquad\quad+\sum_{j=L+1}^{\infty}c_j^2(t)\cdot\left(\frac{1}{4}\nabla\boldsymbol{\sigma}_j(x)g(x)R^{-1}g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_j^{\mathrm{T}}(x)\right),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
&(\mathbf{w}_L(t)-\mathbf{c}_L(t))^{\mathrm{T}}(t_f)\boldsymbol{\sigma}_L(x)\\
&\quad=\hat{e}_L(x,t)+\sum_{j=L+1}^{\infty}c_j(t_f)\sigma_j(x).
\end{aligned}
\tag{22}
$$

Taking the inner product of both sides over Ω, and taking into account the orthonormality of the set $\{\sigma_j\}_1^\infty$, we obtain

$$
\begin{aligned}
&(\dot{\mathbf{w}}_L(t)-\dot{\mathbf{c}}_L(t))+\langle\nabla\boldsymbol{\sigma}_L(x)f(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega^{\mathrm{T}}(\mathbf{w}_L(t)-\mathbf{c}_L(t))\\
&\quad-\left\langle\frac{1}{4}Vec(\nabla\boldsymbol{\sigma}_L(x)g(x)R^{-1}g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_j^{\mathrm{T}}(x)),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega^{\mathrm{T}}\\
&\quad\cdot(\mathbf{w}_L(t)\otimes\mathbf{w}_L(t)-\mathbf{c}_L(t)\otimes\mathbf{c}_L(t))\\
&\qquad=\sum_{j=L+1}^{\infty}c_j(t)\langle\nabla\sigma_j(x)f(x),\boldsymbol{\sigma}_L(x)\rangle_\Omega^{\mathrm{T}}\\
&\qquad\quad+\sum_{j=L+1}^{\infty}c_j^2(t)\left\langle\frac{1}{4}\nabla\sigma_j(x)g(x)R^{-1}g^{\mathrm{T}}(x)\nabla\sigma_j^{\mathrm{T}}(x),\boldsymbol{\sigma}_L(x)\right\rangle_\Omega^{\mathrm{T}}\\
&\qquad\quad\times(\mathbf{w}_L(t)-\mathbf{c}_L(t))(t_f)=0.
\end{aligned}
\tag{23}
$$

Let $A = \langle \nabla \boldsymbol{\sigma}_L(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{\mathrm{T}}$, where A is scalar.

Define $\xi = \mathbf{w}_L(t) - \mathbf{c}_L(t)$, consider the equation

$$\dot{\xi} + A(t) \cdot \xi + f(\xi, t) = 0,$$
$$\xi(t_f) = 0, \tag{24}$$

where

$$f(\xi, t) = -\left\langle \frac{1}{4} Vec(\nabla \boldsymbol{\sigma}_L(x) g(x) R^{-1} g^{\mathrm{T}}(x) \nabla \boldsymbol{\sigma}_L^{\mathrm{T}}(x)), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega}^{\mathrm{T}} \cdot (\mathbf{w}_L(x) \otimes \mathbf{w}_L(x) - \mathbf{c}_L(x) \otimes \mathbf{c}_L(x))$$

is continuously differentiable in a neighborhood of a point (ξ_0, t_0). Since $A(t)$ is also piecewise continuous functions of t, over any finite interval of time $[t_0, t_f]$, the elements of $A(t)$ and $f(\xi, t)$ are bounded. Hence, $\|A(t)\| \leqslant a$, $\|f(\xi, t)\| \leqslant b$ and

$$\|f(x, t) - f(y, t)\| = \|A(t)(x - y)\| \leqslant \|A(t)\| \|x - y\| \leqslant a\|x - y\| \quad \forall x, y \in R^n, \ \forall t \in [t_0, t_f]$$

also

$$\|f(x_0, t)\| = \|A(t) x_0 + f(x, t)\| \leqslant a\|x_0\| + b \leqslant h,$$

for each finite x_0, $\forall t \in [t_0, t_f]$.

Therefore, the system has a unique solution over $[t_0, t_f]$. Since t_f can be arbitrarily large, we can also conclude that if $A(t)$ and $f(x, t)$ are piecewise continuous $\forall t \geqslant t_0$, then the system has a unique solution $\forall t \geqslant t_0$, so (24) can satisfy a local Lipschitz condition (Khalil, 2002).

Noting that

$$\sum_{j=L+1}^{\infty} c_j(t) \langle \nabla \sigma_j(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{\mathrm{T}} + \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} g^{\mathrm{T}}(x) \nabla \sigma_j^{\mathrm{T}}(x), \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega}^{\mathrm{T}}$$

is continuous in t, we invoke the standard result from the theory of ODEs that a continuous perturbation in the system equations and the initial state implies a continuous perturbation of the solution (Arnold, 1973). Note that

$$\left\| \begin{array}{l} \sum_{j=L+1}^{\infty} c_j(t) \langle \nabla \sigma_j(x) f(x), \boldsymbol{\sigma}_L(x) \rangle_{\Omega}^{\mathrm{T}} \\ + \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \begin{pmatrix} \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} \\ \cdot g^{\mathrm{T}}(x) \nabla \sigma_j^{\mathrm{T}}(x) \end{pmatrix}, \boldsymbol{\sigma}_L(x) \right\rangle_{\Omega} \end{array} \right\|_{L_2(\Omega)}$$
$$\leqslant \left\| \sum_{j=L+1}^{\infty} c_j(t) \cdot \langle \nabla \sigma_L(x) f(x), \sigma_L(x) \rangle_{\Omega}^{\mathrm{T}} \right\|_{L_2(\Omega)} + \left\| \sum_{j=L+1}^{\infty} c_j^2(t) \left\langle \begin{array}{l} \frac{1}{4} \nabla \sigma_j(x) g(x) R^{-1} \\ \cdot g^{\mathrm{T}}(x) \nabla \sigma_j^{\mathrm{T}}(x), \boldsymbol{\sigma}_j(x) \end{array} \right\rangle_{\Omega} \right\|_{L_2(\Omega)}$$
$$= \rho(t),$$

here $\rho(t) \to 0$ as L increases.

This implies that for all $\varepsilon > 0$, there exists a $\rho(t) > 0$ such that $\forall t \in [t_0, t_f]$,

$$\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 < \varepsilon. \tag{25}$$

So $\|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2 \to 0$ uniformly in t on Ω as L increases. □

Lemma 3. *Convergence of approximate value function.*

Under the hypotheses of Lemma 1, one has $\|V_L(x, t) - V(x, t)\|_{L_2(\Omega)} \to 0$ uniformly in t on Ω as L increases.

Proof.

$$\begin{aligned} &\|V_L(x, t) - V(x, t)\|_{L_2(\Omega)}^2 \\ &= \int_{\Omega} |V_L(x, t) - V(x, t)|^2 \, \mathrm{d}x \\ &\leqslant \int_{\Omega} |(\mathbf{w}_L(t) - \mathbf{c}_L(t))^{\mathrm{T}} \boldsymbol{\sigma}_L(x)|^2 \, \mathrm{d}x + \int_{\Omega} \left| \sum_{j=L+1}^{\infty} c_j(t) \boldsymbol{\sigma}_j(x) \right|^2 \mathrm{d}x \\ &= (\mathbf{w}_L(t) - \mathbf{c}_L(t))^{\mathrm{T}} \langle \boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L^{\mathrm{T}}(x) \rangle_{\Omega} (\mathbf{w}_L(t) - \mathbf{c}_L(t)) + \int_{\Omega} \left| \sum_{j=L+1}^{\infty} c_j(t) \sigma_j(x) \right|^2 \mathrm{d}x. \end{aligned}$$

By the mean value theorem, $\exists \xi \in \Omega$ such that

$$\|V_L(x, t) - V(x, t)\|_{L_2(\Omega)}^2 = \|\mathbf{w}_L(t) - \mathbf{c}_L(t)\|_2^2 + \lambda(\Omega) \cdot \left| \sum_{j=L+1}^{\infty} c_j(t) \sigma_j(\xi) \right|^2 \to 0$$

uniformly in t on Ω as L increases. □

Lemma 4. *Convergence of value function gradient.*

Under the hypotheses of Lemma 1,

$$\left\| \frac{\partial V_L(x, t)}{\partial x} - \frac{\partial V(x, t)}{\partial x} \right\|_{L_2(\Omega)} \to 0 \text{ uniformly in } t$$

on Ω as L increases.

Proof. From Lemma 2, we have $\|\mathbf{w}_L(t)-\mathbf{c}_L(t)\|_2 \to 0$,

$$\left\|\frac{\partial V_L(x,t)}{\partial x}-\frac{\partial V(x,t)}{\partial x}\right\|^2_{L_2(\Omega)}$$
$$\leqslant \|\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)(\mathbf{w}_L(t)-\mathbf{c}_L(t))\|^2_{L_2(\Omega)} + \left\|\sum_{j=L+1}^{\infty}\nabla\sigma_j^{\mathrm{T}}(x)c_j(t)\right\|^2_{L_2(\Omega)}$$
$$= \|\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)(\mathbf{w}_L(t)-\mathbf{c}_L(t))\|^2_{L_2(\Omega)} + \int_{L_2(\Omega)}\left|\sum_{j=L+1}^{\infty}\nabla\sigma_L^{\mathrm{T}}(x)c_j(t)\right|^2 \mathrm{d}x.$$

By the mean value theorem $\exists\xi\in\Omega$ such that

$$\left\|\frac{\partial V_L(x,t)}{\partial x}-\frac{\partial V(x,t)}{\partial x}\right\|^2_{L_2(\Omega)}$$
$$= \|\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)(\mathbf{w}_L(t)-\mathbf{c}_L(t))\|^2_{L_2(\Omega)} + \lambda(\Omega)\left|\sum_{j=N+1}^{\infty}\nabla\sigma_j^{\mathrm{T}}(x)c_j(t)\right|^2.$$

Since $\nabla\sigma_L^{\mathrm{T}}$ is linear independent and $\|\mathbf{w}_L(t)-\mathbf{c}_L(t)\|_2\to 0$, then $\|\partial V_L(x,t)/\partial x-\partial V(x,t)/\partial x\|_{L_2(\Omega)}\to 0$ uniformly in t on Ω as L increases. □

Lemma 5. *Convergence of control inputs.*

If the conditions of Lemma 1 are satisfied and

$u_L(x,t) = -\frac{1}{2}R^{-1}g^{\mathrm{T}}(x,t)\nabla V_L(x,t)$,

$u(x,t) = -\frac{1}{2}R^{-1}g^{\mathrm{T}}(x,t)\nabla V(x,t)$.

Then $\|u_L(x,t)-u(x,t)\|_{L_2(\Omega)}\to 0$ uniformly in t on Ω as L increases.

Proof.

$$\|u_L(x,t)-u(x,t)\|_{L_2(\Omega)}$$
$$\leqslant \left\|-\frac{1}{2}R^{-1}g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)(\mathbf{w}_L(t)-\mathbf{c}_L(t))\right\|_{L_2(\Omega)} + \left\|\frac{1}{2}\sum_{j=L+1}^{\infty}c_j(t)R^{-1}g^{\mathrm{T}}(t)\nabla\sigma_j(x)\right\|_{L_2(\Omega)}.$$

So $u(x,t) = -\frac{1}{2}\sum_{j=1}^{\infty}c_j(t)R^{-1}g^{\mathrm{T}}(x)\nabla\sigma_j(x)$ implies that the second term on the right-hand side converges to 0. By Lemma 2, Technical Lemmas 2 and 4, we know that

$$\|\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)(\mathbf{w}_L(t)-\mathbf{c}_L(t))\|_{L_2(\Omega)}\to 0.$$

Since $R^{-1}g^{\mathrm{T}}(x,t)$ in continuous on $\Omega\times[t_0,t_f]$ and hence uniformly bounded, we have that

$$\|R^{-1}g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)(\mathbf{w}_L(t)-\mathbf{c}_L(t))\|_{L_2(\Omega)}\to 0. \qquad \square$$

Lemma 6. *Convergence of state trajectory.*

Let $x_L(t)$ be the state using control (12), suppose the hypotheses of Lemma 1 hold. Then $x(t)-x_L(t)\to 0$ uniformly in t on Ω as L increases.

Proof.

$$\dot{x}(t) = f(x)+g(x)u(t) = f(x)-\frac{1}{2}g(x)R^{-1}g^{\mathrm{T}}(x)\frac{\partial V(x,t)}{\partial x},$$
$$\dot{x}_L(t) = f(x_L)+g(x_L)u(t) = f(x_L)-\frac{1}{2}g(x_L)R^{-1}g^{\mathrm{T}}(x_L)\frac{\partial V(x_L,t)}{\partial x_L},$$
$$x_L(t_0)=x(t_0).$$

Since $f(x)-f(x_L)\leqslant L\|x-x_L\|$

$$\dot{x}(t)-\dot{x}_L(t) = f(x)-f(x_L) - \begin{pmatrix}\frac{1}{2}g(x)R^{-1}g^{\mathrm{T}}(x)\frac{\partial V(x,t)}{\partial x}\\ -\frac{1}{2}g(x_L)R^{-1}g^{\mathrm{T}}(x_L)\frac{\partial V_L(x_L,t)}{\partial x_L}\end{pmatrix}$$
$$\leqslant L\|x-x_L\| - \begin{pmatrix}\frac{1}{2}\|R^{-1}\|\cdot(\|g(x)\|_2^2-\|g(x_L)\|_2^2)\frac{\partial V(x,t)}{\partial x}\\ +\frac{1}{2}\left(g(x_L)R^{-1}g^{\mathrm{T}}(x_L)\left(\frac{\partial V(x,t)}{\partial x}-\frac{\partial V(x_L,t)}{\partial x_L}\right)\right)\end{pmatrix}.$$

Define

$$\tilde{x}(t) = x(t)-x_L(t).$$

Consider the equation

$$\dot{\tilde{x}} - L\|\tilde{x}\| + h(\tilde{x},t) = \rho(x), \tag{26}$$
$$\tilde{x}(t_0)=0,$$

where

$$h(\tilde{x},t) = -\left(\frac{1}{2}\|R^{-1}\|\cdot(\|g(x)\|_2^2-\|g(x_L)\|_2^2)\frac{\partial V(x,t)}{\partial x}\right),$$
$$\rho(x) = -\frac{1}{2}\left(g(x_L)R^{-1}g^{\mathrm{T}}(x_L)\left(\frac{\partial V(x,t)}{\partial x}-\frac{\partial V(x_L,t)}{\partial x_L}\right)\right)$$

are continuously differentiable in a neighborhood of a point $(\tilde{x}_0,t_0)$. Over any finite interval of time $[t_0,t_f]$, the elements of $h(\tilde{x},t)$ are bounded. Therefore, (26) has a unique solution. From Lemma 4, $\rho(x)\to 0$ as L increases. We invoke the standard result from the theory of ODEs, as in Lemma 2 proof, so that $\|\tilde{x}\| \to 0$ uniformly in t on Ω as L increases. □

Now, select the Ω_0 in Definition 1 so that $\Omega_0 \subset \Omega$ and $\forall x_0 \in \Omega_0$, $x(t) \in \Omega$, $\forall t \in [t_0, t_f]$. Then, according to Lemma 6, for L large enough, $x_L(t) \in \Omega$, $\forall t \in [t_0, t_f]$. Therefore, the NN approximation property (7) holds to $t \in [t_0, t_f]$.

At this point we have proven uniform convergence in t in the mean of the approximate HJB equation, the NN weights, the approximate value function, the value function gradient and the state trajectory. This demonstrates uniform convergence in t in the mean in Sobolev space $H^{1,2}(\Omega)$. In fact, the next result shows even stronger convergence properties.

Lemma 7. *Uniform convergence.*

Since a local Lipschitz condition holds on (24), then

$$\sup_{x\in\Omega} |V_L(x,t) - V(x,t)| \to 0, \quad \sup_{x\in\Omega} |u_L(x,t) - u(x,t)| \to 0,$$

and

$$\sup_{x\in\Omega} \left| \frac{\partial V_L(x,t)}{\partial x} - \frac{\partial V(x,t)}{\partial x} \right| \to 0.$$

Proof. This follows by noticing that $\|\mathbf{w}_L(t) - c_L(t)\|_2^2 \to 0$ uniformly in t and the series with c_j is uniformly convergent in t, and Hornik et al. (1990). □

The final result shows that if the number L of input-layer units is large enough, the proposed solution method yields an admissible control.

Lemma 8. *Admissibility of* $u_L(x)$.

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

Proof. Define

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

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. □

3.3. Optimal algorithm based on NN approximation

Solving the integration in (15) is expensive computationally. Since evaluation of the L_2 inner product over Ω is required. This can be addressed using the collocation method (Finlayson, 1972). 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 (15) can be rewritten as follows:

$$A = [\boldsymbol{\sigma}_L(x)|_{x_1} \;\; \dots \;\; \boldsymbol{\sigma}_L(x)|_{x_p}]^{\mathrm{T}},$$

$$B = [\boldsymbol{\sigma}_L(x) f(x)|_{x_1} \;\; \dots \;\; \boldsymbol{\sigma}_L(x) f(x)|_{x_p}]^{\mathrm{T}},$$

$$C = \begin{bmatrix} \frac{1}{4}(\nabla\boldsymbol{\sigma}_L(x) g(x) R^{-1} g^{\mathrm{T}}(x) \nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x))|_{x_1} & \dots \\ \frac{1}{4}(\nabla\boldsymbol{\sigma}_L(x) g(x) R^{-1} g^{\mathrm{T}}(x) \nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x))|_{x_p} & \end{bmatrix}^{\mathrm{T}},$$

$$D = [Q(x)|_{x_1} \;\; \dots \;\; Q(x)|_{x_p}]^{\mathrm{T}},$$

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

$$\langle -\dot{\mathbf{w}}_L^{\mathrm{T}}(t)\boldsymbol{\sigma}_L(x), \boldsymbol{\sigma}_L(x)\rangle_\Omega = \lim_{\|\Delta x\|\to 0} -(A^{\mathrm{T}}A)\cdot \dot{\mathbf{w}}_L(t)\cdot \Delta x, \quad (27)$$

$$\langle -\mathbf{w}_L^{\mathrm{T}}(t)\nabla\boldsymbol{\sigma}_L(x) f(x), \boldsymbol{\sigma}_L(x)\rangle_\Omega = \lim_{\|\Delta x\|\to 0} -(A^{\mathrm{T}}B)\cdot \mathbf{w}_L(t)\cdot \Delta x, \quad (28)$$

$$\left\langle \frac{1}{4}\mathbf{w}_L^{\mathrm{T}}(t)\nabla\boldsymbol{\sigma}_L(x) g(x) R^{-1} \cdot g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}\mathbf{w}_L(t), \boldsymbol{\sigma}_L(x)\right\rangle_\Omega = \lim_{\|\Delta x\|\to 0} A^{\mathrm{T}}\mathbf{w}_L^{\mathrm{T}}(t) C \mathbf{w}_L(t)\cdot \Delta x, \quad (29)$$

$$\langle -Q(x), \boldsymbol{\sigma}_L(x)\rangle_\Omega = \lim_{\|\Delta x\|\to 0} -(A^{\mathrm{T}}\cdot D)\cdot \Delta x. \quad (30)$$

This implies that (15) can be converted to

$$\dot{\mathbf{w}}_L(t) = -(A^{\mathrm{T}}A)^{-1}\mathbf{w}_L(t)A^{\mathrm{T}}B + (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{w}_L^{\mathrm{T}}(t) C\mathbf{w}_L(t) - (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}D. \quad (31)$$

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

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

and the nearly optimal control by

$$u_L(t) = -\tfrac{1}{2}R^{-1}g^{\mathrm{T}}(x)\nabla\boldsymbol{\sigma}_L^{\mathrm{T}}(x)\mathbf{w}_L(t). \quad (32)$$

Note that in practice, we use a numerically efficient least-squares relative to solve (31) without matrix inversion.

Remark 2. The closed-loop NN least-squares policy gives correct answer as long as $x \in \Omega$, this control policy would be valid as long as $x(t)$ remains in Ω for all t. This means the set of initial condition Ω, which guarantees that $x(t) \in \Omega$ for all $x(t)$ is smaller than Ω itself. This can be enlarged by carefully selecting larger size of NN.

4. Simulation

We now show the power of our NN control technique for finding nearly optimal fixed-final time controllers to a mobile

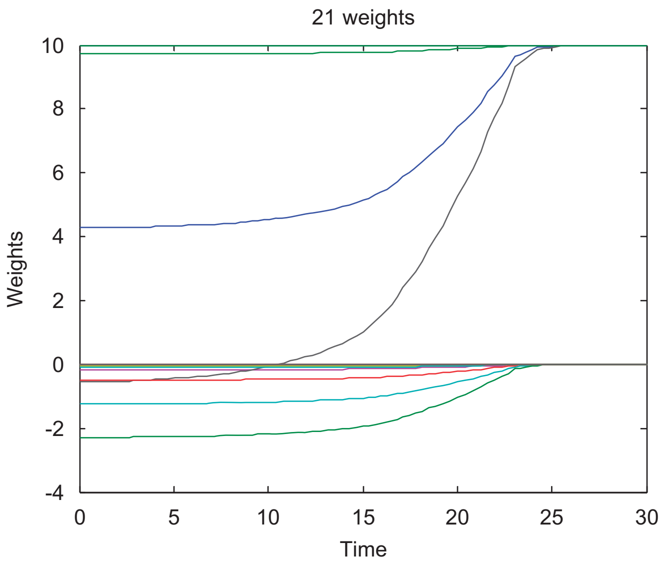


Fig. 1. Nonlinear system weights.

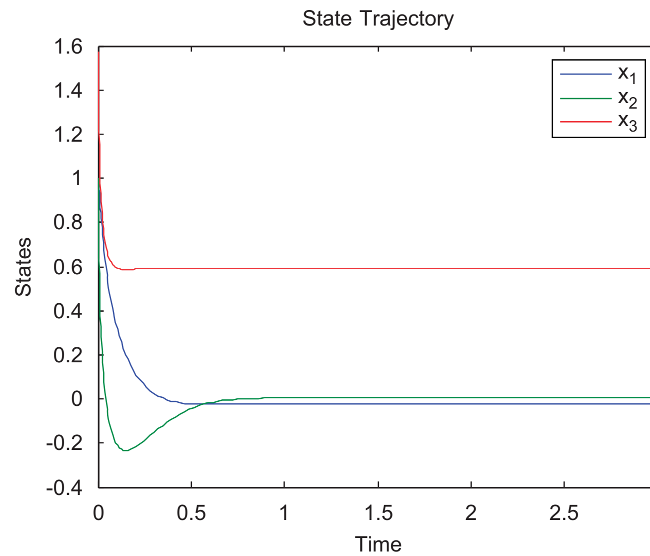


Fig. 2. State trajectory of nonlinear system.

robot, which is a nonholonomic system (Kolmanovsky & McClamroch, 1995). It is known (Brockett, 1983) that there does not exist a continuous time-invariant feedback control law that minimizes the cost. Our method will yield a time-varying gain. From Murray and Sastry, 1991 [32], under some sufficient conditions, a nonholonomic system can be converted into chained form as

$$\dot{x}_1 = u,$$
$$\dot{x}_2 = v,$$
$$\dot{x}_3 = x_1 v. \tag{33}$$

Define performance index

$$V(x(t_0), t_0) = \phi(x(t_f), t_f) + \int_{t_0}^{T} (Q(x) + W(u))\, \mathrm{d}t.$$

Here Q and R are chosen as identity matrices. To solve for the value function of the related optimal control problem, we selected the smooth approximating function

$$\begin{aligned} V(x_1, x_2, x_3) = {} & w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + w_4 x_1 x_2 \\ & + w_5 x_1 x_3 + w_6 x_2 x_3 + w_7 x_1^4 + w_8 x_2^4 \\ & + w_9 x_3^4 + w_{10} x_1^2 x_2^2 + w_{11} x_1^2 x_3^2 + w_{12} x_2^2 x_3^2 \\ & + w_{13} x_1^2 x_2 x_3 + w_{14} x_1 x_2^2 x_3 + w_{15} x_1 x_2 x_3^2 \\ & + w_{16} x_1^3 x_2 + w_{17} x_1^3 x_3 + w_{18} x_1 x_2^3 \\ & + w_{19} x_1 x_3^3 + w_{20} x_2 x_3^3 + w_{21} x_2^3 x_3. \end{aligned} \tag{34}$$

The selection of the NN is usually a natural choice guided by engineering experience and intuition. This is a NN with polynomial activation functions, and hence $V(0) = 0$. This is a power series NN with 21 activation functions containing powers of the state variable of the system upto the fourth order. Convergence was not observed using a NN with only second-order powers of the states. The number of neurons required is chosen

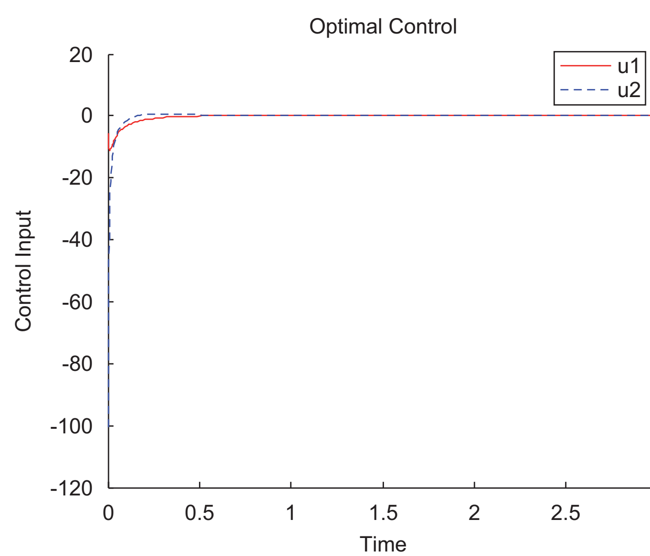


Fig. 3. Optimal NN control law.

to guarantee the uniform convergence of the algorithm. In this example,

$$\mathbf{w}_L(t_f) = [10; 10; 10; 0; 0; 0; 10; 10; 10; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0; 0]$$

and $t_f = 30$ s.

Fig. 1 indicates that weights converge to constant when they are integrated backwards. The time-varying controller (32) is then applied using interpolation. Fig. 2 shows that the systems' states response, including x_1, x_2, and x_3 are all bounded. It can be seen that the states do converge to a value close to the origin. Fig. 3 shows the optimal control converges to zero.

5. 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

Abu-Khalaf, M., & Lewis, F. L. (2005). Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica*, *41*(5), 779–791.

Adams, R., & Fournier, J. (2003). *Sobolev spaces*. (2nd ed.), New York: Academic Press.

Aguilar, L. T., Orlov, Y., & Acho, L. (2003). Nonlinear H_∞-control of nonsmooth time-varying systems with application to friction mechanical manipulators. *Automatica*, *39*, 1531–1542.

Arnold, V. I. (1973). *Ordinary differential equations*. Cambridge, MA: MIT Press.

Bartle, R. G. (1976). *The elements of real analysis*. 2nd ed., New York: Wiley.

Beard, R. (1995). *Improving the closed-loop performance of nonlinear systems*. Ph.D. thesis, Rensselaer Polytechnic Institute, Troy, NY 12180.

Bertsekas, D. P., & Tsitsiklis, J. N. (1995). Neuro-dynamic programming: An overview. *Proceedings of the IEEE conference on decision and control* (pp. 560–564).

Brockett, R. W. (1983). Asymptotic stability and feedback stabilization. In: R. W. Brockett, R. S. Millman, & H. J. Sussmann (Eds.), *Differential geometric control theory* Boston, MA: Birkhauser.

Chen, F. C., & Liu, C. C. (1994). Adaptively controlling nonlinear continuous-time systems using multiplayer neural networks. *IEEE Transactions on Automatic Control*, *39*(6), 1306–1310.

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

Hornik, K., Stinchcombe, M., & White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, *3*, 551–560.

Huang, C. S., Wang, S., & Teo, K. L. (2000). Solving Hamilton–Jacobi–Bellman equations by a modified method of characteristics. *Nonlinear Analysis*, *40*, 279–293.

Huang, J., & Lin, C. F. (1995). Numerical approach to computing nonlinear H_∞ control laws. *Journal of Guidance, Control, and Dynamics*, *18*(5), 989–994.

Isidori, A., & Astolfi, A. (1992). Disturbance attenuation and H_∞-control via measurement feedback in nonlinear systems. *IEEE Transactions on Automatic Control*, *37*(9), 1283–1293.

Khalil, H. K. (2002). *Nonlinear systems*. Upper Saddle River: Prentice-Hall.

Kim, Y. H., Lewis, F. L., & Dawson, D. (2000). Intelligent optimal control of robotic manipulators using neural networks. *Automatica*, *36*, 1355–1364.

Kolmanovsky, I., & McClamroch, N. H. (1995). Developments in nonholonomic control problems. *IEEE Control Systems*, 20–36.

Lewis, F. L., & Syrmos, V. L. (1995). *Optimal control*. New York, NY: Wiley.

Lio, F. D. (2000). On the Bellman equation for infinite horizon problems with unbounded cost functional. *Applied Mathematics and Optimization*, *41*, 171–197.

Miller, W. T., Sutton, R., & Werbos, P. (1990). *Neural networks for control*. Cambridge, Massachusetts: The MIT Press.

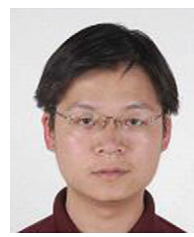
Murray, R.M., Sastry, S.S. (1991). Steering nonholonomic systems in chained form. *Proceedings of the 30th Conference on Decision & Control*, Brighton, England, December 1991.

Parisini, T., & Zoppoli, R. (1998). Neural approximations for infinite-horizon optimal control of nonlinear stochastic systems. *IEEE Transactions on Neural Networks*, *9*(6), 1388–1408.

Sandberg, E. W. (1998). Notes on uniform approximation of time-varying systems on finite time intervals. *IEEE Transactions on Circuits and Systems—1: Fundamental Theory and Applications*, *45*(8),

Saridis, G., & Lee, C. S. (1979). An approximation theory of optimal control for trainable manipulators. *IEEE Transactions on Systems, Man, Cybernetics*, *9*(3), 152–159.

Van der Schaft, A. J. (1992). L_2-gain analysis of nonlinear systems and nonlinear state feedback H_∞ control. *IEEE Transactions on Automatic Control*, *37*(6), 770–784.



Tao Chengwas born in P.R. China in 1976. He received his Bachelor's degree in Electrical Engineering from Hubei Institute of Technology in 1998. He then joined Beijing Polytechnic University from which he received the Master's of Science in Electrical Engineering in 2001. In 2006 he obtained PH.D. degree in Automation and Robotics Research Institute from The University of Texas at Arlington. His research interest is in time-varying optimal nonlinear systems, nonholonomic vehicle systems.

Frank L. Lewis was born in Wurzburg, Germany, subsequently studying in Chile and Gordonstoun School in Scotland. He obtained the Bachelor's Degree in Physics Electrical Engineering and the Master's of Electrical Engineering Degree at Rice University in 1971. In 1977 he received the Master's of Science in Aeronautical Engineering from the University of West Florida. In 1981 he obtained the Ph.D. degree at The Georgia Institute of Technology in Atlanta, where he was employed as a Professor from 1981 to 1990 and is currently an Adjunct Professor. Currently he is a Professor of Electrical Engineering at The University of Texas at Arlington.

Murad Abu-Khalaf was born in Jerusalem, Palestine, in 1977. He joined The University of Texas at Arlington from which he received the Master's of Science in Electrical Engineering in 2000. Currently he is working on his Ph.D. degree at The University of Texas at Arlington. His interest is in the areas of nonlinear control, optimal control and neural network control.